ENT COOPER

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Commissioner
US Department of Commerce
United States Patent and Trademark Office, PCT
2011 South Clark Place Room
CP2/5C24
Arlington, VA 22202
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year)	
30 October 2001 (30.10.01)	
International application No.	**Applicant's or agent's file reference**
PCT/US00/15191	
International filing date (day/month/year)	**Priority date** (day/month/year)
01 June 2000 (01.06.00)	01 June 1999 (01.06.99)
Applicant	
YOUAKIM, Adel et al	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

28 December 2000 (28.12.00)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer Sean Taylor
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference NONE	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/US00/15191	International filing date *(day/month/year)* 01 JUNE 2000	Priority date *(day/month/year)* 01 JUNE 1999
International Patent Classification (IPC) or national classification and IPC Please See Supplemental Sheet.		
Applicant DIGITAL GENE TECHNOLOGIES INC.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 6 sheets.

   [✓] This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 2 sheets.

3. This report contains indications relating to the following items:

   - I [X] Basis of the report
   - II [ ] Priority
   - III [ ] Non-establishment of report with regard to novelty, inventive step or industrial applicability
   - IV [X] Lack of unity of invention
   - V [X] Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
   - VI [ ] Certain documents cited
   - VII [ ] Certain defects in the international application
   - VIII [ ] Certain observations on the international application

Date of submission of the demand	Date of completion of this report
28 DECEMBER 2000	29 AUGUST 2001
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer TARCHAND SAIDHA for Telephone No. (703) 308-0196

Form PCT/IPEA/409 (cover sheet) (July 1998)*

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US00/15191

I. Basis of the report

1. With regard to the **elements** of the international application:*

☐ the international application as originally filed

☒ the description:
pages (See Attached) , as originally filed
pages _____ , filed with the demand
pages _____ , filed with the letter of _____

☒ the claims:
pages (See Attached) , as originally filed
pages _____ , as amended (together with any statement) under Article 19
pages _____ , filed with the demand
pages _____ , filed with the letter of _____

☒ the drawings:
pages (See Attached) , as originally filed
pages _____ , filed with the demand
pages _____ , filed with the letter of _____

☒ the sequence listing part of the description:
pages (See Attached) , as originally filed
pages _____ , filed with the demand
pages _____ , filed with the letter of _____

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language _____ which is:

☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).

☐ the language of publication of the international application (under Rule 48.3(b)).

☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

☒ contained in the international application in printed form.

☒ filed together with the international application in computer readable form.

☒ furnished subsequently to this Authority in written form.

☐ furnished subsequently to this Authority in computer readable form.

☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

☐ The statement that the information recorded in computer readable form is identical to the writen sequence listing has been furnished.

4. ☒ The amendments have resulted in the cancellation of:

☒ the description, pages NONE

☒ the claims, Nos. NONE

☒ the drawings, sheets/~~fig~~ NONE

5. ☐ This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

** Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).*

***Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.*

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US00/15191

IV. Lack of unity of invention

1. In response to the invitation to restrict or pay additional fees the applicant has:

- [ ] restricted the claims.
- [ ] paid additional fees.
- [ ] paid additional fees under protest.
- [ ] neither restricted nor paid additional fees.

2. [X] This Authority found that the requirement of unity of invention is not complied with and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

- [ ] complied with.
- [X] not complied with for the following reasons:

Please See Supplemental Sheet.

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

- [X] all parts.
- [ ] the parts relating to claims Nos. _.

Supplemental Box
(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: Boxes I - VIII

CLASSIFICATION:
The International Patent Classification (IPC) and/or the National classification are as listed below:
IPC(7): C12Q 1/00, 1/68; C12P 21/06; C12N 1/20, 15/00; A61K 38/00; C07K 1/00, 16/00; C07H 21/02, 21/04 and US Cl.: 435/4, 6, 69.1, 252.3, 320.1; 530/300, 350, 387.1; 536/23.1, 23.4, 23.5

I. BASIS OF REPORT:

This report has been drawn on the basis of the description,
page(s) 1-50, 52-75, as originally filed.
page(s) NONE, filed with the demand.
and additional amendments:
Page 51, filed with the letter of 03 August 2001.

This report has been drawn on the basis of the claims,
page(s) 76-79, as originally filed.
page(s) NONE, as amended under Article 19.
page(s) NONE, filed with the demand.
and additional amendments:
NONE

This report has been drawn on the basis of the drawings,
page(s) 1-4, as originally filed.
page(s) NONE, filed with the demand.
and additional amendments:
NONE

This report has been drawn on the basis of the sequence listing part of the description:
page(s) 1-8, 10-25, as originally filed.
pages(s) NONE, filed with the demand.
and additional amendments:
Page 9, filed with the letter of 03 August 2001.

IV. LACK OF UNITY OF INVENTION:

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2, and 13.3 is not complied with for the following reasons:
This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1.

Group I, claim(s)1-15, 19-20, 24-27 & 32, drawn to polynucleotide of SEQ ID NO : 1-10 and the encoded polypeptides, vectors, host cells and related method and kit.

Group II, claim(s)1-15, 19-20, 24-27 & 32, drawn to polynucleotide of SEQ ID NO : 11-20 and the encoded polypeptides, vectors, host cells and related method and kit.

Group III, claim(s)1-15, 19-20, 24-27 & 32, drawn to polynucleotide of SEQ ID NO : 21-30 and the encoded polypeptides, vectors, host cells and related method and kit.

Group IV, claim(s)1-15, 19-20, 24-27 & 32, drawn to polynucleotide of SEQ ID NO : 31-40 and the encoded polypeptides, vectors, host cells and related method and kit.

Group V, claim(s)1-15, 19-20, 24-27 & 32, drawn to polynucleotide of SEQ ID NO : 41-43 and the encoded polypeptides, vectors, host cells and related method and kit.

Group VI, claim(s) 16-18 & 28-31, drawn to antibody of polypeptide of SEQ ID NO : 1-10.

Supplemental B x
(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: Boxes I - VIII

Group VII, claim(s) 16-18 & 28-31, drawn to antibody of polypeptide of SEQ ID NO : 11-20.

Group VIII, claim(s) 16-18 & 28-31, drawn to antibody of polypeptide of SEQ ID NO : 21-30.

Group IX, claim(s) 16-18 & 28-31, drawn to antibody of polypeptide of SEQ ID NO : 31-40.

Group X, claim(s) 16-18 & 28-31, drawn to antibody of polypeptide of SEQ ID NO : 41-43.

Group XI, claim(s) 21-23, drawn to methods of diagnosing, identifying, preventing and treating using the polypeptides or polynucleotides of SEQ ID NOs: 1-10.

Group XII, claim(s) 21-23, drawn to methods of diagnosing, identifying, preventing and treating using the polypeptides or polynucleotides of SEQ ID NOs: 11-20.

Group XIII, claim(s) 21-23, drawn to methods of diagnosing, identifying, preventing and treating using the polypeptides or polynucleotides of SEQ ID NOs: 21-30.

Group XIV, claim(s) 21-23, drawn to methods of diagnosing, identifying, preventing and treating using the polypeptides or polynucleotides of SEQ ID NOs: 31-40.

Group XV, claim(s) 21-23, drawn to methods of diagnosing, identifying, preventing and treating using the polypeptides or polynucleotides of SEQ ID NOs: 41-43.

Unity of Invention - Nucleotide sequences

The Commissioner has decided sua sponte to partially waive 37 CFR 1.475 and 1.499 et seq. to permit applicants to claims up to ten (10) nucleotide sequences which do not have the same or corresponding special technical feature, without the payment of an additional fee. The PCT permits inventions which lack unity of invention to be maintained in the same international application for payment of additional fees. Thus, in international applications, for each group for which applicant has paid additional international search and/or preliminary examination fees, the PTO has determined that up to four (4) such additional sequences per group is a reasonable number for examination. Further, claims directed to the selected sequences will be examined with claims drawn to any sequence combinations which have a common technical feature with the selected sequences. Nucleotide sequence encoding the same protein are considered to satisfy the unity of invention standard and will continue to be examined together.
and it considers that the International Application does not comply with the requirements of unity of invention (Rules 13.1, 13.2 and 13.3) for the reasons indicated below:

The inventions listed as Groups I-XV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Groups I-V have a special technical feature of the nucleotide sequence encoding the polypeptide (see note below), which Groups VI-X and XI-XV do not share. Groups VI-X have a special technical feature of the antibody which Groups I-V and XI-XV do not share. Groups XI-XV have a special technical feature of diagnosis, treatment and prevention, which Groups I-V and Groups VI-X do not share. Thus the various groups discussed above show a lack of unity of invention.

The Commissioner has decided sua sponte to partially waive 37 CFR 1.475 and 1.499 et seq. to permit applicants to claims up to ten (10) nucleotide sequences which do not have the same or corresponding special technical feature, without the payment of an additional fee. The PCT permits inventions which lack unity of invention to be maintained in the same international application for payment of additional fees. Thus, in international applications, for each group for which applicant has paid additional international search and/or preliminary examination fees, the PTO has determined that up to four (4) such additional sequences per group is a reasonable number for examination. Further, claims directed to the selected sequences will be examined with claims drawn to any sequence combinations which have a common technical feature with the selected sequences. Nucleotide sequence encoding the same protein are considered to satisfy the unity of invention standard and will continue to be examined together.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US00/15191

V. Reason d statement under Articl 35(2) with regard t novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. **statement**

Novelty (N)	Claims	1-2, 5, 7-32	YES
	Claims	3, 4, 6	NO
Inventive Step (IS)	Claims	1-2, 5, 7-32	YES
	Claims	3, 4, 6	NO
Industrial Applicability (IA)	Claims	1-32	YES
	Claims	NONE	NO

2. **citations and explanations (Rule 70.7)**

Claims 3, 4, 6 lack novelty under PCT Article 33(2) as being anticipated by Foss et al. [BBA 1402 (1998): 17-28]. Foss et al. teach MECL1 cDNA (Accession No : 13640) and the gene (see Figure 1) sequence which is 99.4% homologous to Applicants' SEQ ID NO : 15. Foss's SEQ ID NO : 15 meets each of the limitations of claim 3, reciting '95% identical'; claim 4, reciting 'at least 10 bases in length'; and claim 6, reciting 'encoding a fragment of the polypeptide'.

Claims 1-2, 5 & 7-32 meet the criteria set out in PCT Article 33(2)-(3), because the prior art does not teach or fairly suggest the claimed polypeptide or polynucleotide sequences, vectors, host cells, method of making the polypeptides, antibodies to the polypeptides and the various methods of use of the polypeptides or polynucleotides.

Claims 1-32 have industrial applicability under PCT Article 33(4), because the subject matter claimed can be made or used in industry.

NEW CITATIONS

NONE

<400> 43
taatggaatg tctcggtgta aagctgacca ttcccattcg ttctattctg agataggaga 60
aaaccgccct gtggctggag gtgagatatg ctggcagcaa tactgctctg ttactccttg 120
ctacactgag atgtttgggt aaagagaaac ataaatctag cctacgtgca catctgggca 180
cagtaccttt ccttgaactt attcgtgata cagattcctt tgctcacatg tttccctgct 240
gaccttcttc ccacctgttg ccctgctaca ctcccctcgc taagacagta aaaataatga 300
tcaataaata ctgagggaac tcagaggcca gcgccggtgc gggtcctcca catgctgagc 360
gccggtcctg ggcccactgt t 381

<210> 44
<211> 49
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence:synthetic primer

<400> 44
gaattcaact ggaagcggcc cgcaggaatt ttttttttt ttttttvnn 49

<210> 45
<211> 16
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer

<400> 45
aggtcgacgg tatcgg 16

<210> 46
<211> 16
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence:synthetic primer

<400> 46
ggtcgacggt atcggn 16

<210> 47
<211> 15
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence:synthetic primer

<400> 47
gagctccacc gcggt 15

<210> 48
<211> 16
<212> DNA
<213> Artificial Sequence

Sutcliffe, J.G., et al *Proc Natl Acad Sci U S A* 2000 Feb 29; 97(5):1976-1981, International published application PCT/US99/23655, U.S. Patent No. 5,459,037, U.S. Patent No. 5,807,680, and U.S. Patent No. 6,030,784, hereby incorporated herein by reference. Preferably, prior to the application of the TOGA method or other methods, the isolated RNA was enriched to form a starting polyA-containing mRNA population by methods known in the art. In such a preferred embodiment, the TOGA method further comprises an additional Polymerase Chain Reaction ("PCR") step performed using one of four 5' PCR primers and cDNA templates prepared from a population of antisense complimentary RNA ("cRNAs"). A final PCR step using one of a possible 256 5' PCR primers and a universal 3' PCR primer produced as PCR products, cDNA fragments that corresponded to a 3'-region of the starting mRNA population. The produced PCR products were then identified by a) the sequence of at least the 5' seven base pairs, preferably the sequence of the entire fragment, and b) the length of the fragment. These two parameters, sequence and fragment length, were used to compare the obtained PCR products to a database of known polynucleotide sequences.

The method yields Digital Sequence Tags (DSTs), that is, polynucleotides that are expressed sequence tags (ESTs) of the 3' end of mRNAs. DSTs that showed changes in relative levels during barrier breakdown were selected for further study. The intensities of the laser-induced fluorescence of the labeled PCR products were compared across samples isolated at different time intervals after treatment.

In general, double-stranded cDNA is generated from poly(A)-enriched cytoplasmic RNA extracted from the tissue samples of interest using an equimolar mixture of all 48 5'-biotinylated anchor primers of a set to initiate reverse transcription. One such suitable set is G-A-A-T-T-C-A-A-C-T-G-G-A-A-G-C-G-G-C-C-C-G-C-A-G-G-A-A-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-V-N-N (SEQ ID NO: 44), where V is A, C or G and N is A, C, G or T. One member of this mixture of 48 anchor primers initiates synthesis at a fixed position at the 3' end of all copies of each mRNA species in the sample, thereby defining a 3' endpoint for each species, resulting in biotinylated double stranded cDNA.

Each biotinylated double stranded cDNA sample was cleaved with the restriction endonuclease MspI, which recognizes the sequence CCGG. The 3' fragments of cDNA

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 December 2000 (07.12.2000)

PCT

(10) International Publication Number
WO 00/73324 A2

(51) International Patent Classification[7]: C07K

(21) International Application Number: PCT/US00/15191

(22) International Filing Date: 1 June 2000 (01.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/137,058 1 June 1999 (01.06.1999) US

(71) Applicant *(for all designated States except US)*: **DIGITAL GENE TECHNOLOGIES INC.** [US/US]; 11149 North Torrey Pines Road, Suite 110, La Jolla, CA 92037 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **YOUAKIM, Adel** [CA/US]; 7304 175th Street SW, Edmonds, WA 98026 (US). **DUBOSE, Robert, Finley** [US/US]; 6151 156th Place Southeast, Bellevue, WA 98006 (US). **SIMS, John, E.** [US/US]; 4207 43rd Avenue NE, Seattle, WA 98105 (US). **PRIBYL, Thomas, M.** [US/US]; 2609 Pico Place, Apartment 127, San Diego, CA 92109 (US). **HILLBUSH, Brian, S.** [US/US]; 13058 Caminito Angelico, San Diego, CA 92130 (US). **HASEL, Karl, W.** [CH/US]; 247 South Nardo Avenue, Solana Beach, CA 92075 (US).

(74) Agent: **ZIMMERMAN, Roger, P.**; McDonnell Boehnen Hulbert & Berghoff, 300 South Wacker Drive, 32nd Floor, Chicago, IL 60606 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/73324 A2

(54) Title: MODULATION OF GENE EXPRESSION IN GASTROINTESTINAL INFLAMMATION

(57) Abstract: Polynucleotides, polypeptides, kits and methods are provided related to regulated genes characteristic of gastrointestinal inflammation.

MODULATION OF GENE EXPRESSION IN GASTROINTESTINAL INFLAMMATION (Case No. 98,501-A)

REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application S.N. 60/137,058, filed June 1, 1999, which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

Damage to the intestinal epithelial barrier is a hallmark of inflammatory bowel diseases (IBD). Studying the factors that influence the integrity of the intestinal epithelial barrier *in vivo* is a difficult task for a number of reasons, including the complexity of the tissue itself (there are numerous cell types in the gut including epithelial, stromal, endocrine, neuronal and hematopoietic) and the technical problems associated with tissue manipulation in animals or in isolated organs. As a result of these reasons, a number of *in vitro* models of intestinal epithelial barrier function have been developed over the years.

The best characterized of these models is the T84 intestinal epithelial barrier system (Dharmsathaphorn, K., J.A. McRoberts, K.G. Mandel, L.D. Tisdale, and H. Masui. 1984. A human colonic tumor cell line that maintains vectorial electrolyte transport. *Am J. Physiol.* 246- G204-0208; Madara, J.L., and K. Dharmsathaphorn. 1985. Occluding junction structure-function relationships in a cultured epithelial monolayer. *J. Cell Biol.* 101. 2124-2133). Although T84 cells were derived from a human colonic adenocarcinoma, they have retained many of the properties associated with normal colonic crypt cells. T84 cells form polarized monolayers that exhibit high electrical resistance as well as vectorial fluid and chloride secretion reminiscent of colonic crypts *in vivo*. These properties are directly dependent on a complex of proteins referred to as tight junctions and T84 cells express many of the known members of this complex. These cells also respond to proinflammatory cytokines, such as interferon-gamma (IFN-γ), by decreasing barrier function (Madara, J.L., and J. Stafford. 1989. Interferon-gamma directly affects barrier function of cultured intestinal epithelial monolayers. *J. Clin. Invest.* 83: 724-727; Adams, R.B., S.M. Planchon, and J.K. Roche. 1993. IFN-gamma modulation of epithelial

barrier function. Time course, reversibility, and site of cytokine binding. *J. Immunol.* 150: 2356-2363.), and by up-regulating MHC Class II molecules and antigen presenting activity (Hershberg, R.M., D.H. Cho, A. Youakim, M.B. Bradley, J.S. Lee, P.E. Framson and G.T. Nepom. 1998. Highly polarized HLA class II antigen processing and presentation by human intestinal epithelial cells. *J. Clin. Invest.* 102: 792-803).

SUMMARY OF THE INVENTION

The T84 cell model system has been used in studies to examine how the epithelial barrier is regulated by various agents such as interferon-gamma and cells of the immune system (i.e., neutrophils). Such studies have examined the mechanism of barrier breakdown and recovery in response to such agents and to identify proteins and genes that may prevent barrier breakdown or stimulate barrier recovery. Molecules have been identified that correspond to genes that are regulated by the treatment of T84 cells with proinflammatory cytokines, such as interferon-gamma. Such molecules are useful in therapeutic and diagnostic applications in the treatment of IBD and other gut pathologies.

The present invention provides novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polynucleotides and the polypeptides. Also provided are diagnostic methods for detecting disorders related to the polypeptides and the polynucleotides encoding them, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

In one embodiment, the invention provides an isolated nucleic acid molecule comprising a polynucleotide chosen from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34,

SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43. Another embodiment comprises an isolated nucleic acid molecule at least 95% identical to the isolated nucleic acid molecule of SEQ ID NO:1-43. A further embodiment comprises an isolated nucleic acid molecule at least ten bases in length that is hybridizable to the isolated nucleic acid molecule of SEQ ID NO:1-43 under stringent conditions.

In another embodiment, the invention provides an isolated polypeptide encoded by a polynucleotide chosen from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43. In another embodiment, the invention provides an isolated nucleic acid molecule encoding the polypeptide of the present invention.

In a further embodiment, the invention provides a substantially pure isolated DNA molecule suitable for use as a probe for genes regulated in gastrointestinal inflammation, chosen from the group consisting of the DNA molecules identified in Table 1, having a 5' partial nucleotide sequence and length as described by their digital address, and having a characteristic regulation pattern in T84 cells treated with interferon-gamma.

The present invention also provides a system and method for detecting the presence of a gene regulated in gastrointestinal inflammation. In one embodiment, the present invention provides a kit for suitable for detecting the presence of a gene regulated in gastrointestinal inflammation, comprising at least one polynucleotide of the present invention, or fragment thereof having at least 10 contiguous bases, in an amount sufficient for at least one assay; label means; instructions for use; and suitable packaging material. In one embodiment, the polynucleotide is chosen from the group consisting of SEQ ID

NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43. Another embodiment comprises a polynucleotide at least 95% identical to the isolated nucleic acid molecule of SEQ ID NO:1-43. A further embodiment comprises a polynucleotide at least ten bases in length that is hybridizable to the isolated nucleic acid molecule of SEQ ID NO:1-43 under stringent conditions. In yet another embodiment, the polynucleotide is chosen from the group consisting of the DNA molecules identified in Table 1, having a 5' partial nucleotide sequence and length as described by their digital address, and having a characteristic regulation pattern in T84 cells treated with interferon-gamma.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other features, aspects, and advantages of the present invention will become better understood with reference to the following description, appended claims, and accompanying drawings where:

Figure 1 is a graphical representation of the results of TOGA analysis (TOtal Gene expression Analysis) using a 5' PCR primer with parsing bases GGCC, showing PCR products produced from mRNA extracted from untreated T84 cells (Figure 1A), T84 cells 4 hours after treatment with interferon-gamma (Figure 1B), 24 hours after treatment with interferon-gamma (Figure 1C), and T84 cells 44 hours after treatment with interferon-gamma (Figure 1D), where the vertical index line indicates a PCR product of about 374 b.p. present in T84 cells that is enriched in treated T84 cells, where the ordinate is in arbitrary units of fluorescence intensity and the abscissa is length of PCR product in nucleotides;

Figure 2 is a graphical representation of more detailed analysis of the 374 b.p. PCR product indicated in Figure 1; Figure 2A shows the PCR product obtained using an extended 5' primer as described in the text; Figure 2B shows the PCR products obtained using the original PCR primers, and in Figure 2C, the traces from Figure 2A and 2B are overlaid, demonstrating that the PCR product of the isolated and sequenced clone is the same length as the original PCR product, where the ordinate is in arbitrary units of fluorescence intensity and the abscissa is length of PCR product in nucleotides; and

Figure 3 is a graphical representation of the results of Northern Blot analysis of clone IMX 43, SEQ ID NO: 6, where an agarose gel containing poly A enriched mRNA from the four T84 experimental samples (control, 4 hour, 24 hour and 44 hour) as well as size standards was blotted after electrophoresis, the Northern blot probed with radiolabelled IMX 43, imaged using a phosphorimager and quantified.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a

"mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:1-43. For example, the polynucleotide can contain all or part of the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:1-43, or the complement thereof, or the cDNA. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 microgram/milliliter denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH_2PO_4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are

well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formulation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS – STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about ten-fold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.).

The translated amino acid sequence, beginning with the methionine, is identified although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by the translation of these alternative open reading frames are specifically contemplated by the present invention.

SEQ ID NO:1-43 and the translations of SEQ ID NO: 1-43 are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from the translations of SEQ ID NO:1-43 may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

The present invention also relates to the genes corresponding to SEQ ID NO:1-43, and translations of SEQ ID NO:1-43. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologues. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. Therefore, from a deduced amino acid sequence, a signal sequence and mature sequence can be identified.

Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

"Identity" per se has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, (1988); BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, (1993); COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, (1994); SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, (1987); and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, (1991).) While there exists a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans. (Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988).) Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in "Guide to Huge Computers," Martin J. Bishop, ed., Academic Press, San Diego, (1994), and Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988). Methods for aligning polynucleotides or polypeptides are codified in computer programs, including the GCG program package (Devereux, J., et al., Nucleic Acids Research (1984) 12(l):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F. et al., J. Molec. Biol. 215:403 (1990), Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711 (using the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981).)

When using any of the sequence alignment programs to determine whether a particular sequence is, for instance, 95% identical to a reference sequence, the parameters are set so that the percentage of identity is calculated over the full length of the reference polynucleotide and that gaps in identity of up to 5% of the total number of nucleotides in the reference polynucleotide are allowed.

A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245 (1990).) The term "sequence" includes nucleotide and amino acid sequences. In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred

parameters used in a FASTDB search of a DNA sequence to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=l, Joining Penalty=30, Randomization Group Length=0, and Cutoff Score=l, Gap Penalty=5, Gap Size Penalty 0.05, and Window Size=500 or query sequence length in nucleotide bases, whichever is shorter. Preferred parameters employed to calculate percent identity and similarity of an amino acid alignment are: Matrix=PAM 150, k-tuple=2, Mismatch Penalty= 1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=l, Gap Penalty=5, Gap Size Penalty=0.05, and Window Size=500 or query sequence length in amino acid residues, whichever is shorter.

As an illustration, a polynucleotide having a nucleotide sequence of at least 95% "identity" to a sequence contained in SEQ ID NO:1-43 means that the polynucleotide is identical to a sequence contained in SEQ ID NO:1-43 or the cDNA except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the total length (not just within a given 100 nucleotide stretch). In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to SEQ ID NO:1-43, up to 5% of the nucleotides in the sequence contained in SEQ ID NO:1-43 or the cDNA can be deleted, inserted, or substituted with other nucleotides. These changes may occur anywhere throughout the polynucleotide.

Further embodiments of the present invention include polynucleotides having at least 80% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to a sequence contained in SEQ ID NO:1-43. Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the polynucleotides having at least 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity will encode a polypeptide identical to an amino acid sequence contained in the translations of SEQ ID NO:1-43.

Similarly, by a polypeptide having an amino acid sequence having at least, for example, 95% "identity" to a reference polypeptide, is intended that the amino acid sequence of the polypeptide is identical to the reference polypeptide except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the total length of the reference polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be

deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

Further embodiments of the present invention include polypeptides having at least 80% identity, more preferably at least 85% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to an amino acid sequence contained in translations of SEQ ID NO: 1-43. Preferably, the above polypeptides should exhibit at least one biological activity of the protein.

In a preferred embodiment, polypeptides of the present invention include polypeptides having at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98%, or 99% similarity to an amino acid sequence contained in translations of SEQ ID NO:1-43.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993) reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL -1a. They used random mutagenesis to generate over 3,500 individual IL- 1 a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See Gayle et al., (1993), Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein

the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to

increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in that shown in SEQ ID NO:1-43. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in that shown in SEQ ID NO:1-43. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, and more nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, to the end of SEQ ID NO:1-43. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in the translations of SEQ ID NO:1-43. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a

part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, or 61 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50 or 60, amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of the translations of SEQ ID NO:1-43 falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of

the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Additional embodiments include chimeric antibodies, e.g., humanized versions of murine monoclonal antibodies. Such humanized antibodies may be prepared by known techniques, and offer the advantage of reduced immunogenicity when the antibodies are administered to humans. In one embodiment, a humanized monoclonal antibody comprises the variable region of a murine antibody (or just the antigen binding site thereof) and a constant region derived from a human antibody. Alternatively, a humanized antibody fragment may comprise the antigen binding site of a murine monoclonal antibody and a variable region fragment (lacking the antigen-binding site) derived from a human antibody. Procedures for the production of chimeric and further engineered monoclonal antibodies include those described in Riechmann et al. (*Nature 332*:323, 1988), Liu et al. (*PNAS 84*:3439, 1987), Larrick et al. (*Bio/Technology 7*:934, 1989), and Winter and Harris (*TIPS 14*:139, May, 1993).

One method for producing an antibody comprises immunizing a non-human animal, such as a transgenic mouse, with a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-43, whereby antibodies directed against the polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-43 are generated in said animal. Procedures have been developed for generating human antibodies in non-human animals. The antibodies may be partially human, or preferably completely human. Non-human animals (such as transgenic mice) into which genetic material encoding one or more human immunoglobulin chains has been introduced may be employed. Such transgenic mice may be genetically altered in a variety of ways. The genetic manipulation may result in human immunoglobulin polypeptide chains replacing endogenous immunoglobulin chains in at least some (preferably virtually all) antibodies produced by the animal upon immunization. Antibodies produced by immunizing transgenic animals with a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-43 are provided herein.

Mice in which one or more endogenous immunoglobulin genes are inactivated by various means have been prepared. Human immunoglobulin genes have been introduced into the mice to replace the inactivated mouse genes. Antibodies produced in the animals incorporate human immunoglobulin polypeptide chains encoded by the human genetic

material introduced into the animal. Examples of techniques for production and use of such transgenic animals are described in U.S. Patents 5,814,318, 5,569,825, and 5,545,806, which are incorporated by reference herein.

Monoclonal antibodies may be produced by conventional procedures, e.g., by immortalizing spleen cells harvested from the transgenic animal after completion of the immunization schedule. The spleen cells may be fused with myeloma cells to produce hybridomas, by conventional procedures.

A method for producing a hybridoma cell line comprises immunizing such a transgenic animal with a immunogen comprising at least seven contiguous amino acid residues of a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-43; harvesting spleen cells from the immunized animal; fusing the harvested spleen cells to a myeloma cell line, thereby generating hybridoma cells; and identifying a hybridoma cell line that produces a monoclonal antibody that binds a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-43. Such hybridoma cell lines, and monoclonal antibodies produced therefrom, are encompassed by the present invention. Monoclonal antibodies secreted by the hybridoma cell line are purified by conventional techniques.

Antibodies may be employed in an *in vitro* procedure, or administered *in vivo* to inhibit biological activity induced by a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-43. Disorders caused or exacerbated (directly or indirectly) by the interaction of such polypeptides of the present invention with cell surface receptors thus may be treated. A therapeutic method involves *in vivo* administration of a blocking antibody to a mammal in an amount effective for reducing a biological activity induced by a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-43.

Also provided herein are conjugates comprising a detectable (e.g., diagnostic) or therapeutic agent, attached to an antibody directed against a polypeptide translated from a nucleotide sequence chosen from SEQ ID NO:1-43. Examples of such agents are well known, and include but are not limited to diagnostic radionuclides, therapeutic

radionuclides, and cytotoxic drugs. The conjugates find use in *in vitro* or *in vivo* procedures.

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric

secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-0 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0 232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).) Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for

propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, 293, Bowes melanoma cells and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, PNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated

transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat

polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:1-43. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:1-43 will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian

Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library).) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. The polynucleotides of SEQ ID NO:1-43 can be used for this analysis of individual humans.

First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and

the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an

identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class H HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (^{125}I, ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^{3}H), indium (^{112}In), and technetium (^{99m}Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of

protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ^{131}I, ^{112}In, ^{99m}Tc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of ^{99m}Tc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activitiy

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and

proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoinununne disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation,

Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis),

Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsielia, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A

polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteocarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by

attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2), Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli.* Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled

competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (c) determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color,

skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, circadian rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by activin or inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 80%, preferably at least 85%, more preferably at least 90%, most preferably at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:1-43.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:1-43 in the range of positions beginning with the nucleotide at about the position of the 5' nucleotide of the clone sequence and ending with the nucleotide at about the position of the 3' nucleotide of the clone sequence.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:1-43 in the range of positions beginning with the nucleotide at about the position of the 5' nucleotide of the start codon and ending with the nucleotide at about the position of the 3' nucleotide of the clone sequence as defined for SEQ ID NO:1-43.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:1-43 in the range of positions beginning with the nucleotide at about the position of the 5' nucleotide of the first amino acid of the signal peptide and ending with the nucleotide at about the position of the 3' nucleotide of the clone sequence as defined for SEQ ID NO:1-43.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:1-43.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:1-43.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:1-43 beginning with the nucleotide at about the position of the 5' nucleotide of the first amino acid of the signal peptide and ending with the nucleotide at about the position of the 3' nucleotide of the clone sequence as defined for SEQ ID NO:1-43.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:1-43.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:1-43, which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample

with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:1-43.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of nucleotide sequences of SEQ ID NO:1-43.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence

selected from the group consisting of: a nucleotide sequence of SEQ ID NO:1-43. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in an amino acid sequence translated from SEQ ID NO:1-43.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in an amino acid sequence translated from SEQ ID NO:1-43, in the range of positions beginning with the residue at about the position of the first amino acid of the secreted portion and ending with the residue at about the last amino acid of the open reading frame.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in an amino acid sequence translated from SEQ ID NO:1-43.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in an amino acid sequence translated from SEQ ID NO:1-43.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to acids in an amino acid sequence translated from SEQ ID NO:1-43.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-43, which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected

from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-43.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-43.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-43.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide

wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-43.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encodes a polypeptide comprising an amino acid sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-43.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of amino acid sequences translated from SEQ ID NO:1-43. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

The present invention also includes a diagnostic system, preferably in kit form, for assaying for the presence of the polypeptide of the present invention in a body sample, such brain tissue, cell suspensions or tissue sections; or a body fluid sample, such as CSF,

blood, plasma or serum, where it is desirable to detect the presence, and preferably the amount, of the polypeptide of this invention in the sample according to the diagnostic methods described herein.

In a related embodiment, a nucleic acid molecule can be used as a probe (i.e., an oligonucleotide) to detect the presence of a polynucleotide of the present invention, a gene corresponding to a polynucleotide of the present invention, or a mRNA in a cell that is diagnostic for the presence or expression of a polypeptide of the present invention in the cell. The nucleic acid molecule probes can be of a variety of lengths from at least about 10, suitably about 10 to about 5000 nucleotides long, although they will typically be about 20 to 500 nucleotides in length. Hybridization methods are extremely well known in the art and will not be described further here.

In a related embodiment, detection of genes corresponding to the polynucleotides of the present invention can be conducted by primer extension reactions such as the polymerase chain reaction (PCR). To that end, PCR primers are utilized in pairs, as is well known, based on the nucleotide sequence of the gene to be detected. Preferably, the nucleotide sequence is a portion of the nucleotide sequence of a polynucleotide of the present invention. Particularly preferred PCR primers can be derived from any portion of a DNA sequence encoding a polypeptide of the present invention, but are preferentially from regions which are not conserved in other cellular proteins.

Preferred PCR primer pairs useful for detecting the genes corresponding to the polynucleotides of the present invention and expression of these genes are described in the Examples, including the corresponding Tables. Nucleotide primers from the corresponding region of the polypeptides of the present invention described herein are readily prepared and used as PCR primers for detection of the presence or expression of the corresponding gene in any of a variety of tissues.

The diagnostic system includes, in an amount sufficient to perform at least one assay, a subject polypeptide of the present invention, a subject antibody or monoclonal antibody, and/or a subject nucleic acid molecule probe of the present invention, as a separately packaged reagent.

In another embodiment, a diagnostic system, preferably in kit form, is contemplated for assaying for the presence of the polypeptide of the present invention or an antibody immunoreactive with the polypeptide of the present invention in a body fluid sample. Such diagnostic kit would be useful for monitoring the fate of a therapeutically administered polypeptide of the present invention or an antibody immunoreactive with the polypeptide of the present invention. The system includes, in an amount sufficient for at least one assay, a polypeptide of the present invention and/or a subject antibody as a separately packaged immunochemical reagent.

In another embodiment, a diagnostic system, preferably in kit form, is contemplated for assaying for the presence of a nucleic acid in a body fluid or tissue sample that is characteristic of inflammatory bowel disease, using a polynucleotide of the present invention. The system includes, in an amount sufficient for at least one assay, at least one polynucleotide of the present invention or fragments thereof as well as other reagents.

Instructions for use of the packaged reagent(s) are also typically included.

As used herein, the term "package" refers to a solid matrix or material such as glass, plastic (e.g., polyethylene, polypropylene or polycarbonate), paper, foil and the like capable of holding within fixed limits a polypeptide, polyclonal antibody, or monoclonal antibody of the present invention. Thus, for example, a package can be a glass vial used to contain milligram quantities of a contemplated polypeptide or antibody or it can be a microtiter plate well to which microgram quantities of a contemplated polypeptide or antibody have been operatively affixed (i.e., linked) so as to be capable of being immunologically bound by an antibody or antigen, respectively.

"Instructions for use" typically include a tangible expression describing the reagent concentration or at least one assay method parameter, such as the relative amounts of reagent and sample to be admixed, maintenance time periods for reagent/ sample admixtures, temperature, buffer conditions and the like.

A diagnostic system of the present invention preferably also includes a label or indicating means capable of signaling the formation of an immunocomplex containing a polypeptide or antibody molecule of the present invention.

The word "complex" as used herein refers to the product of a specific binding reaction such as an antibody-antigen or receptor-ligand reaction. Exemplary complexes are immunoreaction products.

As used herein, the terms "label means" and "indicating means" in their various grammatical forms refer to single atoms and molecules that are either directly or indirectly involved in the production of a detectable signal to indicate the presence of a complex. Any label or indicating means can be linked to or incorporated in a polynucleotide, an expressed protein, polypeptide, or antibody molecule that is part of an antibody or monoclonal antibody composition of the present invention or used separately, and those atoms or molecules can be used alone or in conjunction with additional reagents. Such labels are themselves well-known in clinical diagnostic chemistry and constitute a part of this invention only insofar as they are utilized with otherwise novel proteins methods and/or systems. Suitable label means comprise enzymes, radioisotopes, fluorescent compounds, colloidal metals, chemiluminescent compounds, phosphorescent compounds, or bioluminescent compounds.

The labeling means can be a fluorescent labeling agent that chemically binds to polynucleotides, antibodies or antigens, without preventing interaction with the binding partner, to form a fluorochrome (dye) that is a useful immunofluorescent tracer. Suitable fluorescent labeling agents are fluorochromes such as fluorescein isocyanate (FIC), 6-FAM, fluorescein isothiocyante (FITC), 5-dimethylamine-1-naphthalenesulfonyl chloride (DANSC), tetramethylrhodamine isothiocyanate (TRITC), lissamine, rhodamine 8200 sulphonyl chloride (RB 200 SC) and the like. A description of immunofluorescence analysis techniques is found in DeLuca, "Immunofluorescence Analysis", in *Antibody As a Tool*, Marchalonis et al., Eds., John Wiley & Sons, Ltd., pp. 189-231 (1982), which is incorporated herein by reference. Other suitable labeling agents are known to those skilled in the art.

In preferred embodiments, the indicating group is an enzyme, such as horseradish peroxidase (HRP), glucose oxidase, or the like. In such cases where the principal indicating group is an enzyme such as HRP or glucose oxidase, additional reagents are required to visualize the fact that a receptor-ligand complex (immunoreactant) has formed. Such additional reagents for HRP include hydrogen peroxide and an oxidation dye precursor such as diaminobenzidine. An additional reagent useful with glucose oxidase is 2,2'-amino-di-(3-ethyl-benzthiazoline-G-sulfonic acid) (ABTS).

Radioactive elements are also useful labeling agents and are used illustratively herein. An exemplary radiolabeling agent is a radioactive element that produces gamma ray emissions. Elements which themselves emit gamma rays, such as ^{124}I, ^{125}I, ^{128}I, ^{132}I and ^{51}Cr represent one class of gamma ray emission-producing radioactive element indicating groups. Particularly preferred is ^{125}I. Another group of useful labeling means are those elements such as ^{11}C, ^{18}F, ^{15}O and ^{13}N which themselves emit positrons. The positrons so emitted produce gamma rays upon encounters with electrons present in the animal's body. Also useful is a beta emitter, such ^{111} indium or ^{3}H.

The linking of labels or labeling of polypeptides and proteins is well known in the art. For instance, antibody molecules produced by a hybridoma can be labeled by metabolic incorporation of radioisotope-containing amino acids provided as a component in the culture medium (see, e.g., Galfre et al., *Meth. Enzymol.*, 73:3-46 (1981)). The techniques of protein conjugation or coupling through activated functional groups are particularly applicable (see, for example, Aurameas, et al., *Scand. J. Immunol.*, Vol. 8 Suppl. 7:7-23 (1978); Rodwell et al., *Biotech.*, 3:889-894 (1984); and U.S. Pat. No. 4,493,795).

The diagnostic systems can also include, preferably as a separate package, a specific binding agent. A "specific binding agent" is a molecular entity capable of selectively binding a reagent species of the present invention or a complex containing such a species, but is not itself a polypeptide or antibody molecule composition of the present invention. Exemplary specific binding agents are second antibody molecules, complement proteins or fragments thereof, S. aureus protein A, and the like. Preferably the specific binding agent binds the reagent species when that species is present as part of a complex.

In preferred embodiments, the specific binding agent is labeled. However, when the diagnostic system includes a specific binding agent that is not labeled, the agent is typically used as an amplifying means or reagent. In these embodiments, the labeled specific binding agent is capable of specifically binding the amplifying means when the amplifying means is bound to a reagent species-containing complex.

The diagnostic kits of the present invention can be used in an "ELISA" format to detect the quantity of the polypeptide of the present invention in a sample. "ELISA" refers to an enzyme-linked immunosorbent assay that employs an antibody or antigen bound to a solid phase and an enzyme-antigen or enzyme-antibody conjugate to detect and quantify the amount of an antigen present in a sample. A description of the ELISA technique is found in Sites et al., *Basic and Clinical Immunology,* 4th Ed., Chap. 22, Lange Medical Publications, Los Altos, CA (1982) and in U.S. Patent No. 3,654,090; Patent No. 3,850,752; and Patent No. 4,016,043, which are all incorporated herein by reference.

Thus, in some embodiments, a polypeptide of the present invention, an antibody or a monoclonal antibody of the present invention can be affixed to a solid matrix to form a solid support that comprises a package in the subject diagnostic systems.

A reagent is typically affixed to a solid matrix by adsorption from an aqueous medium, although other modes of affixation applicable to polynucleotides, proteins and polypeptides can be used that are well known to those skilled in the art. Exemplary adsorption methods are described herein.

Useful solid matrices are also well known in the art. Such materials are water insoluble and include the cross-linked dextran available under the trademark SEPHADEX from Pharmacia Fine Chemicals (Piscataway, NJ), agarose; polystyrene beads of ut 1 micron (μm) to about 5 millimeters (mm) in diameter available from several suppliers (e.g., Abbott Laboratories, Chicago, IL), polyvinyl chloride, polystyrene, cross-linked polyacrylamide, nitrocellulose- or nylon-based webs (sheets, strips or paddles) or tubes, plates or the wells of a microtiter plate, such as those made from polystyrene or polyvinylchloride.

The reagent species, labeled specific binding agent, or amplifying reagent of any diagnostic system described herein can be provided in solution, as a liquid dispersion or as a substantially dry power, e.g., in lyophilized form. Where the indicating means is an enzyme, the enzyme's substrate can also be provided in a separate package of a system. A solid support such as the before-described microtiter plate and one or more buffers can also be included as separately packaged elements in this diagnostic assay system.

The packaging materials discussed herein in relation to diagnostic systems are those customarily utilized in diagnostic systems.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

EXAMPLE 1

Identification and Characterization of Polynucleotides Regulated by Interferon-γ in T84 cells

T84 cells were plated on 75 mm polycarbonate transwell filter inserts (Costar) and grown in DME/F12 (1:1) containing 10% heat-inactivated bovine calf serum. The cells were maintained at confluence for 2-3 days and integrity of the epithelial barrier was determined by measuring transepithelial electrical resistance (TER) using an EVOM epithelial voltohmmeter (World Precision Instruments). When the TER values were greater than 1000 ohms/cm^2 and were stable, cells were treated with interferon-γ (30 ng/ml, Genzyme) added to the basolateral side of the membrane. At various times after treatment (4, 24 and 44h), TERs were measured to monitor the interferon-induced disruption of the barrier and RNA was harvested from the cells at those time points using TRIzol reagent (Life Technologies). The isolated RNA was subsequently used for TOGA analysis.

Isolated RNA was analyzed using a method of simultaneous sequence-specific identification of mRNAs known as TOGA (TOtal Gene expression Analysis) described in

Sutcliffe, J.G., et al *Proc Natl Acad Sci U S A* 2000 Feb 29; 97(5):1976-1981, International published application PCT/US99/23655, U.S. Patent No. 5,459,037, U.S. Patent No. 5,807,680, and U.S. Patent No. 6,030,784, hereby incorporated herein by reference. Preferably, prior to the application of the TOGA method or other methods, the isolated RNA was enriched to form a starting polyA-containing mRNA population by methods known in the art. In such a preferred embodiment, the TOGA method further comprises an additional Polymerase Chain Reaction ("PCR") step performed using one of four 5' PCR primers and cDNA templates prepared from a population of antisense complimentary RNA ("cRNAs"). A final PCR step using one of a possible 256 5' PCR primers and a universal 3' PCR primer produced as PCR products, cDNA fragments that corresponded to a 3'-region of the starting mRNA population. The produced PCR products were then identified by a) the sequence of at least the 5' seven base pairs, preferably the sequence of the entire fragment, and b) the length of the fragment. These two parameters, sequence and fragment length, were used to compare the obtained PCR products to a database of known polynucleotide sequences.

The method yields Digital Sequence Tags (DSTs), that is, polynucleotides that are expressed sequence tags (ESTs) of the 3' end of mRNAs. DSTs that showed changes in relative levels during barrier breakdown were selected for further study. The intensities of the laser-induced fluorescence of the labeled PCR products were compared across samples isolated at different time intervals after treatment.

In general, double-stranded cDNA is generated from poly(A)-enriched cytoplasmic RNA extracted from the tissue samples of interest using an equimolar mixture of all 48 5'-biotinylated anchor primers of a set to initiate reverse transcription. One such suitable set is G-A-A-T-T-C-A-A-C-T-G-G-A-A-G-C-G-G-C-C-C-G-C-A-G-G-A-A-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-T-V-N-N (SEQ ID NO: 44), where V is A, C or G and N is A, C, G or T. One member of this mixture of 48 anchor primers initiates synthesis at a fixed position at the 3' end of all copies of each mRNA species in the sample, thereby defining a 3' endpoint for each species, resulting in biotinylated double stranded cDNA.

Each biotinylated double stranded cDNA sample was cleaved with the restriction endonuclease MspI, which recognizes the sequence CCGG. The 3' fragments of cDNA

were then isolated by capture of the biotinylated cDNA fragments on a streptavidin-coated substrate. Suitable streptavidin-coated substrates include microtitre plates, PCR tubes, polystyrene beads, paramagnetic polymer beads and paramagnetic porous glass particles. A preferred streptavidin-coated substrate is a suspension of paramagnetic polymer beads (Dynal, Inc., Lake Success, NY).

After washing the streptavidin-coated substrate and captured biotinylated cDNA fragments, the cDNA fragment product was released by digestion with NotI, which cleaves at an 8-nucleotide sequence within the anchor primers but rarely within the mRNA-derived portion of the cDNAs. The 3' MspI-NotI fragments, which are of uniform length for each mRNA species, were directionally ligated into ClaI-, NotI-cleaved plasmid pBC SK^+ (Stratagene, La Jolla, CA) in an antisense orientation with respect to the vector's T3 promoter, and the product used to transform Escherichia coli SURE cells (Stratagene). The ligation regenerates the NotI site, but not the MspI site. Each library contained in excess of 5×10^5 recombinants to ensure a high likelihood that the 3' ends of all mRNAs with concentrations of 0.001% or greater were multiply represented. Plasmid preps (Qiagen) were made from the cDNA library of each sample under study.

An aliquot of each library was digested with MspI, which effects linearization by cleavage at several sites within the parent vector while leaving the 3' cDNA inserts and their flanking sequences, including the T3 promoter, intact. The product was incubated with T3 RNA polymerase (MEGAscript kit, Ambion) to generate antisense cRNA transcripts of the cloned inserts containing known vector sequences abutting the MspI and NotI sites from the original cDNAs.

At this stage, each of the cRNA preparations was processed in a three-step fashion. In step one, 250 ng of cRNA was converted to first-strand cDNA using the 5' RT primer (A-G-G-T-C-G-A-C-G-G-T-A-T-C-G-G, (SEQ ID NO: 45). In step two, 400 pg of cDNA product was used as PCR template in four separate reactions with each of the four 5' PCR primers of the form G-G-T-C-G-A-C-G-G-T-A-T-C-G-G-N (SEQ ID NO: 46), each paired with a "universal" 3' PCR primer G-A-G-C-T-C-C-A-C-C-G-C-G-G-T (SEQ ID NO: 47).

In step three, the product of each subpool was further divided into 64 subsubpools (2ng in 20µl) for the second PCR reaction, with 100 ng each of the fluoresceinated "universal" 3' PCR primer, the oligonucleotide (SEQ ID NO: 47) conjugated to 6-FAM and the appropriate 5' PCR primer of the form C-G-A-C-G-G-T-A-T-C-G-G-N-N-N-N (SEQ ID NO: 48), using a program that included an annealing step at a temperature X slightly above the T_m of each 5' PCR primer to minimize artifactual mispriming and promote high fidelity copying. Each polymerase chain reaction step was performed in the presence of TaqStart antibody (Clonetech).

The products from the final polymerase chain reaction step for each of the tissue samples were resolved on a series of denaturing DNA sequencing gels using the automated ABI Prizm 377 sequencer. Data were collected using the GeneScan software package (ABI) and normalized for amplitude and migration. Complete execution of this series of reactions generated 64 product subpools for each of the four pools established by the 5' PCR primers of the first PCR reaction, for a total of 256 product subpools for the entire 5' PCR primer set of the second PCR reaction.

The mRNA samples from each timepoint after induction of barrier breakdown as described above were analyzed. Table 1 is a summary of the expression levels of 79 mRNAs determined from cDNA. These cDNA molecules are identified by their digital address, that is, a partial 5' terminus nucleotide sequence coupled with the length of the molecule, as well as the relative amount of the molecule produced at different time intervals after treatment. The 5' terminus partial nucleotide sequence is determined by the recognition site for MspI and the nucleotide sequence of the parsing bases of the 5' PCR primer used in the final PCR step. The digital address length of the fragment was determined by interpolation on a standard curve and, as such, may vary $\pm$ 1-2 b.p. from the actual length as determined by sequencing.

For example, the entry in Table 1 that describes a DNA molecule identified by the digital address MspI GTGC 257, is further characterized as having a 5' terminus partial nucleotide sequence of CGGTGC and a digital address length of 257 b.p. The DNA molecule identified as MspI GTGC 257 is further described as being expressed at increasing levels after the induction of barrier breakdown. However, the treatment results

in a marked decrease of the expression of MspI ACTG 191, producing a 13-fold decrease between the untreated control cells and the induced cells 44 hours after induction.

Similarly, the other 77 DNA molecules identified in Table 1 by their MspI digital addresses are further characterized by 1) level of gene expression in untreated T84 cells, 2) level of gene expression 4 hours after barrier breakdown induction, 3) level of gene expression 24 hours after barrier breakdown induction and 4) level of gene expression 44 hours after barrier breakdown induction. Many of the isolated clones were further characterized in Table 2 and their nucleotide sequences are provided as SEQ ID NO: 1-43 in the Sequence Listing below.

The data shown in Figure 1 were generated with a 5'-PCR primer (C-G-A-C-G-G-T-A-T-C-G-G-G-G-C-C, SEQ ID NO: 49) paired with the "universal" 3' primer (SEQ ID NO: 47) labeled with 6-carboxyfluorescein (6FAM, ABI) at the 5' terminus. PCR reaction products were resolved by gel electrophoresis on 4.5% acrylamide gels and fluorescence data acquired on ABI377 automated sequencers. Data were analyzed using GeneScan software (Perkin-Elmer).

The results of TOGA analysis using a 5' PCR primer with parsing bases GGCC (SEQ ID NO: 49) are shown in Figure 1, which presents the results of TOGA analysis using a 5' PCR primer with parsing bases GGCC, showing PCR products produced from mRNA extracted from untreated T84 cells (Figure 1A), T84 cells 4 hours after treatment with interferon-gamma (Figure 1B), 24 hours after treatment with interferon-gamma (Figure 1C), and T84 cells 44 hours after treatment with interferon-gamma (Figure 1D). The vertical index line indicates a PCR product of about 374 b.p. that is present in T84 cells and whose expression increases after induction of barrier breakdown in T84 cells.

In order to verify that the clones isolated are from the same peak, PCR primers were designed based on the determined sequence and PCR was performed using the cDNA produced in the first PCR reaction as substrate. For example, for the 374 b.p. product disclosed above, oligonucleotides were synthesized corresponding to the 5' PCR primer in the second PCR step extended at the 3' end with additional nucleotides from the

sequences adjacent to the terminal MspI sites in the GenBank sequences. This primer had the sequence: GAT CGA ATC CGG GGC CCA GGC CTA TGC CTC (SEQ ID NO: 50).

The products were separated by electrophoresis and the length of the clone was compared to the length of the original PCR product as shown in Figure 2. In Figure 2, the upper panel shows the length (as peak position) of the PCR product derived from the isolated clone as described above. Figure 2A shows the PCR product obtained using an extended 5' PCR primer. Figure 2B shows the PCR products obtained using the original PCR primers SEQ ID NO: 49 and SEQ ID NO: 47 (compare to the top panel in Figure 1). In Figure 2C. the traces from Figure 2A and 2B are overlaid, demonstrating that the PCR product of the isolated and sequenced clone is the same length as the original PCR product.

Table 4, below, shows the results of the verification of candidate matches to database entries. IMX 2 and 53 were cloned and assayed as described above for GGCC 374 to verify these clones. In each case, oligonucleotides were synthesized corresponding to the 5' PCR primer in the second PCR step extended at the 3' end with additional nucleotides from the sequences adjacent to the terminal MspI sites in the identified corresponding GenBank sequences.

IMX 2 (SEQ ID NO:4) and IMX 53 (SEQ ID NO:37) were cloned and assayed as described above for GGCC 374 to verify that the clones isolated are from the TOGA peaks (GAAT 140 and TGCC 265, respectively). The 5' oligonucleotides used in the above mentioned assay, were synthesized which contain a portion of the MspI site (CGG) and the parsing bases (GAAT or TGCC, respectively) with clone sequence 3' to the parsing bases. The sequences of these oligonucleotides are listed in Table 4.

The same procedure was also used to verify candidate matches to database entries. As indicated in Table 4, oligonucleotides were synthesized that extend at the 3' end with additional nucleotides from the sequences adjacent to the terminal MspI sites in the identified corresponding GenBank sequences.

Other PCR products, notably IMX 22 (SEQ ID NO: 1) and IMX 38 (SEQ ID NO: 2), were also differentially represented, and appeared to migrate in positions that indicate that the PCR products were novel based on comparison to data extracted from GenBank. In these cases, the PCR product was isolated, cloned into a TOPO vector (Invitrogen) and sequenced on both strands.

Figure 3 is a graphical representation of the results of Northern Blot analysis of clone IMX 43, SEQ ID NO: 6, where an agarose gel containing poly A enriched mRNA from the four T84 experimental samples (control, 4 hour, 24 hour and 44 hour) as well as size standards was blotted after electrophoresis and probed with radiolabelled IMX 43, SEQ ID NO: 6, imaged using a phosphorimager and quantified.

TABLE 1

Seq ID	Clone ID	Digital Address (Msp1)	Control	4 hour	24 hour	44 hour
1st Tier Increases						
3	IMX 1	GAAT 131	161	1327	2081	1952
4	IMX 2	GAAT 140	152	834	1437	1395
	IMX 3	GAAT 172	47	165	305	347
5	IMX 5	AGTC 136	91	113	522	714
32	IMX 6	GCTT 88	123	861	878	597
33	IMX 7	GCCT 97	278	339	1312	1672
8	IMX 8	ACAG 154	289	575	1579	1380
9	IMX 9	AAAT 80	78	64	331	304
10	IMX 11	TGAA 134	150	400	507	492
11	IMX 12	CTTT 91	233	664	707	248
14	IMX 16	TGAA 160	549	1525	3261	2035
18	IMX 20	GTGC 257	35	99	599	768
19	IMX 21A	CTTC 364	35	60	330	749
1	IMX 22	TTCA 420	44	36	484	757
20	IMX 23	ATAT 318	61	715	997	124
21	IMX 24	AGCA 261	62	44	354	598
22	IMX 25	GCCC 379	36	36	246	225
23	IMX 26	TCCG 333	137	113	297	804
24	IMX 27	CCGT 349	283	1484	364	568
25	IMX 29	TTAC 201	99	72	293	463
29	IMX 37	ACTA 116	318	968	1712	629
6	IMX 43	GGCC 374	1391	2111	5940	5314
7	IMX 45	CGAT 346	59	96	128	299
2	IMX 48	CCTT 346	19	67	77	78
36	IMX 50	TCCA 371	111	422	624	464
	IMX 51	AGCT 341	548	3319	1724	3218
		AGAG 319	131	775	90	916
1st Tier Decreases						
12	IMX 13	ACTG 191	1315	360	367	102
13	IMX 14	AGCG 149	904	843	448	164
17	IMX 19	CCCG 250	3915	3590	1202	981
26	IMX 34	CTCC 257	218	84	52	37
27	IMX 35	CCAC 303	114	102	76	26

TABLE 1, Continued

Seq ID	Clone ID	Digital Address (Msp1)	Control	4 hour	24 hour	44 hour
2nd Tier Increases						
30	IMX 38	TGTG 164	1477	3740	5478	6947
31	IMX 41	TCCT 85	822	2852	1711	2291
37	IMX 53	TGCC 265	561	820	795	2344
38	IMX 54	CACT 344	70	69	228	203
		CTTA 284	146	109	573	575
		CGAA 345	105	151	145	333
		ATTC 369	31	23	32	93
		CTTA 345	41	85	122	99
2nd Tier Decreases						
28	IMX 36	AAGA 356	528	435	151	98
34	IMX 46	CGGG 230	1514	751	1368	369
35	IMX 47	CCAG 260	221	60	64	77
		ACCC 253	584	581	368	106
		CGTC 223	569	573	212	153
		TCCT 325	498	517	248	137
		CGGA 234	212	150	71	67
		ATCC 459	79	71	45	25
		CAGC 453	109	42	58	46
3rd Tier Increases						
39	IMX 58	GCGG 164	1523	3889	4854	5084
	IMX 61	CGCT 264	965	3060	2757	3181
		AGCC 135	1737	2354	4085	5227
		CGAC 345	32	52	95	79
		AGGG 246	362	619	598	1062
		TTCC 85	1321	2400	2835	3797
		CGTG 140	53	71	147	140
		TACA 91	2158	3366	3313	5640
3rd Tier Decreases						
		AGCG 101	1137	1364	675	311
		CGGC 97	629	314	411	187
		ACCA 105	1041	741	998	334
		ACTC 409	302	197	122	111
		TGTG 251	470	435	302	176
		CCGC 382	179	177	94	68

TABLE 1, Continued

Seq ID	Clone ID	Digital Address (Msp1)	Control	4 hour	24 hour	44 hour
Medium Annotations						
15	IMX 17	GGGC 306	294	613	1128	1247
	IMX 59	GTAG 454	66	248	260	169
		TGGA 391	21	43	357	390
		CGGA 438	97	301	879	946
		GATG 205	969	893	1523	3850
		AATG 248	660	1032	1619	1751
		ATCT 382	159	127	70	58
		GCAG 392	246	136	103	85
		CAGG 89	588	519	196	150
		ATTT 136	128	75	30	32
		GGCG 364	224	265	120	50
		AAGA 432	458	477	118	72
High Confidence Annotations						
16	IMX 18	CACC 336	1133	3363	3930	5050
	IMX 57	GACT 96	2088	3689	5075	6520
		GGCC 374	795	1531	4918	5196
		AATC 334	142	411	235	161

TABLE 2

Seq ID	Clone ID	Digital Address (Msp1)	Database Match (Accession #)	Relative DST Amount*				Validation Method
				0 hour	4 hour	24 hour	44 hour	
3	IMX 1	GAAT 131	Homo sapiens K-12 protein precursor mRNA, complete cds. (U77643)	60	322	288	633	R
5	IMX 5	AGTC 136	Human 1-8U gene from interferon-inducible gene family (X57352)	104	251	839	1239	R
33	IMX 7	GCCT 97	Human mRNA for protease activator hPA28 subunit beta, complete cds (D45248/HUMPHPA28A)	66	142	422	456	R
8	IMX 8	ACAG 154	EST Soares NhHMPu S1 Homo sapiens cDNA clone (AA252172)	60	225	423	517	R
10	IMX 11	TGAA 134	Human LMP7 gene, exons 1-7 (L11045)	655	1070	1049	1382	R
12	IMX 13	ACTG 191	EST Soares total fetus Nb2HF8 9w Homo Sapiens cDNA clone 784238 3' (AA446881)	1122	520	518	618	R
13	IMX 14	AGCG 149	H. sapiens CpG island DNA genomic Mse1 fragment, clone 43f4, reverse read cpg43f4.rtla (Z65513)	887	916	818	1246	R
14	IMX 16	TGAA 160	H. sapiens genes TAP1, TAP2, LMP2, LMP7 and DOB (X66401)	161	524	823	662	R
15	IMX 17	GGGC 306	Homo sapiens mRNA for proteosome subunit MECl-1 (Y13640)	353	786	1111	933	R
16	IMX 18	CACC 336	H. sapiens mRNA for metallothionein isoform 2 (X97260)	247	712	940	950	R
17	IMX 19	CCCG 250	Human creatine kinase isoenzyme CK-B gene, exon 8 (M21243)	1269	1377	890	665	R
18	IMX 20	GTGC 257	Human mRNA for complement component C1r (X04701)	761	1304	1709	2120	R
19	IMX 21A	CTTC 364	EST nn44f07.s1 NCI_CGAP_GC5 Homo sapiens cDNA clone IMAGE:1086757 similar to contains L1.t3 L1 repetitive element; mRNA sequence (AA582172)	-0-	287	594	668	R
1	IMX 22	TTCA 420	NOVEL	351	376	214	378	R
20	IMX 23	ATAT 318	Homo sapiens CLN3 gene, complete CDS (X99832)	395	789	532	400	R
21	IMX 24	AGCA 261	Human complement component C4A gene, exons 10 through 41 (M59815/HUMC4AA2)	0	65	249	297	R
22	IMX 25	GCCC 379	EST Homo sapiens cDNA clone 122225 3' similar to contains MER22 repetitive element (T98760)	247	454	263	260	R
23	IMX 26	TCCG 333	Homo sapiens transcription regulator protein (BACH 1) (AF026199)	547	1159	1258	2116	R
24	IMX 27	CCGT 349	GSP Z93241 Human DNA sequence	79	105	106	114	NBA
25	IMX 29	TTAC 201	Human low density lipoprotein receptor gene, exon 18 (L00352)	452	600	784	1190	R
26	IMX 34	CTCC 257	Human Mitochondrial DNA control region (M76299)	873	978	902	906	R
27	IMX 35	CCAC 303	Human HepG2 3' region Mbol cDNA, clone hmd3a03m3 (D17187)	735	1056	783	675	R

GSP = Genome Sequencing Project, EST = Expressed Sequence Tag, NA = Not Applicable, R = Reverse Transcriptase-Polymerase Chain Reaction, NBA = Northern Blot Analysis

*Relative DST Amount = Relative DST Amount for Northern blots have been corrected for background and normalized to the hybridization signal of cyclophilin. The values were obtained by exposing the Northern blot to a phosphorimaging screen and quantitated using the Phosphorimager SI.

TABLE 2, Continued

Seq ID	Clone ID	Digital Address (MspI)	Database Match (Accession #)	Relative DST Amount				Validation Method
				0 hour	4 hour	24 hour	44 hour	
28	IMX 36	AAGA 356	Human liver fatty acid binding protein (L-FABP) mRNA, complete cds (M10617)	1115	794	466	322	R
29	IMX 37	ACTA 116	H. sapiens mitochondrial genome (X62996)	112	629	569	471	R
30	IMX 38	TGTG 164	H. sapiens gene TAP1, TAP 2, LMP2, LMP7 and DOB (X66401)	579	1366	1295	1947	R
6	IMX 43	GGCC 374	H. sapiens mRNA for interferon-induced 17kDa membrane protein (X84958)	202	495	1815	2417	NBA
7	IMX 45	CGAT 346	EST Jia bone marrow stroms Homo sapiens cDNA clone 1091412 3' similar to WP:CE03924 BETA TRANSDUCIN (AA599195)	325	862	785	869	R
34	IMX 46	CGGG 230	EST Homo sapiens cDNA clone IMAGE: 1117847 (AA618529)	596	706	475	523	R
35	IMX 47	CCAG 260	EST Strategene ovarian cancer (#937219) Homo sapiens cDNA clone 595329 3' (AA173988)	517	694	443	217	R
2	IMX 48	CCTT 346	NOVEL	0	147	348	701	R
36	IMX 50	TCCA 371	EST qu13e05.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMA mRNA sequence (AI283282)	347	1127	1159	1027	R
38	IMX 54	CACT 344	Human variant urokinase plasinogen activator receptor (uPAR2) mRNA, partial cds (U09347)	156	307	391	425	R
39	IMX 58	GCGG 164	H. sapiens mRNA for leukocyte antigen CD97 (X84700)	107	309	586	932	R

GSP = Genome Sequencing Project, EST = Expressed Sequence Tag, NA = Not Applicable, R = Reverse Transcriptase-Polymerase Chain Reaction, NBA = Northern Blot Analysis

*Relative DST Amount = Relative DST Amount for Northern blots have been corrected for background and normalized to the hybridization signal of cyclophilin. The values were obtained by exposing the Northern blot to a phosphorimaging screen and quantitated using the Phosphorimager SI.

TABLE 3

Seq ID	Clone ID	Digital Address (Msp1)	Database Match (Accession #)	% Homology	Nucleotide homology	
					DST nucleotide range (bp #)	Database nucleotide range (bp #)
3	IMX 1	GAAT 131	Homo sapiens K-12 protein precursor mRNA, complete cds. (U77643)	97	1 – 78	1645 – 1722
4	IMX 2	GAAT 140	Homo sapiens K-12 protein precursor mRNA, complete cds. (U77643)	94	1 – 78	1645 – 1722
5	IMX 5	AGTC 136	Human 1-8U gene from interferon-inducible gene family (X57352)	97	1 – 79	722 – 800
32	IMX 6	GCTT 88	Qk33b09.x1 NCI_CGAP_Co8 H. sapiens cDNA clone IMAGE:1870745 3' similar to SW:SYS_HUMAN P49591 SERYL-TRNA SYNTHETASE (AI 245918)	95	1 – 37	13 – 49
33	IMX 7	GCCT 97	Human mRNA for protease activator hPA28 subunit beta, complete cds (D45248/HUMPHPA28A)	97	1 – 37	789 – 825
8	IMX 8	ACAG 154	EST Soares NhHMPu S1 Homo sapiens cDNA clone (AA252172)	99	1 – 98	2 – 99
9	IMX 9	AAAT 80	Human interferon-induced 17-kDa protein mRNA, complete cds (M13755)	100	1 – 21	608 – 628
10	IMX 11	TGAA 134	Human LMP7 gene, exons 1-7 (L11045)	99	1 – 85	3805 – 3889
11	IMX 12	CTTT 91	EST Soares parathyroid tumor NbHPA H. sapiens cDNA clone 1360348 3', mRNA sequence (AA781258)	98	1 – 40	3 – 42
12	IMX 13	ACTG 191	EST Soares total fetus Nb2HF8 9w Homo Sapiens cDNA clone 784238 3' (AA446881)	99	1 – 138	32 – 169
13	IMX 14	AGCG 149	H. sapiens CpG island DNA genomic Mse1 fragment, clone 43f4, reverse read cpg43f4.rtla (Z65513)	99	1 – 94	63 – 156
14	IMX 16	TGAA 160	H. sapiens genes TAP1, TAP2, LMP2, LMP7 and DOB (X66401)	98	1 – 104	37840 – 37943
15	IMX 17	GGGC 306	Homo sapiens mRNA for proteosome subunit MECl-1 (Y13640)	99	1 – 255	712 – 966
16	IMX 18	CACC 336	H. sapiens mRNA for metallothionein isoform 2 (X97260)	99	1 – 280	48 – 326
17	IMX 19	CCCG 250	Human creatine kinase isoenzyme CK-B gene, exon 8 (M21243)	99	1 – 193	224 – 416
18	IMX 20	GTGC 257	Human mRNA for complement component C1r (X04701)	99	1 – 202	2085 – 2287
19	IMX 21A	CTTC 364	EST nn44f07.s1 NCI_CGAP_GC5 Homo sapiens cDNA clone IMAGE:1086757 similar to contains L1.t3 L1 repetitive element; mRNA sequence (AA582172)	99	1 – 308	3 – 310
1	IMX 22	TTCA 420	NOVEL	NA	NA	NA
20	IMX 23	ATAT 318	Homo sapiens CLN3 gene, complete CDS (X99832)	97	1 – 265	12484 – 12749
21	IMX 24	AGCA 261	Human complement component C4A gene, exons 10 through 41 (M59815/HUMC4AA2)	99	1 – 207	11501 – 11707
22	IMX 25	GCCC 379	EST Homo sapiens cDNA clone 122225 3' similar to contains MER22 repetitive element (T98760)	98	56 – 310	1 – 256
23	IMX 26	TCCG 333	Homo sapiens transcription regulator protein (BACH 1) (AF026199)	99	1 – 283	388 – 670
24	IMX 27	CCGT 349	GSP Z93241 Human DNA sequence	99	1 – 297	106912 – 107533

GSP = Genome Sequencing Project, EST = Expressed Sequence Tag, NA = Not Applicable

TABLE 3, Continued

Seq ID	Clone ID	Digital Address (MspI)	Database Match (Accession #)	% Homology	Nucleotide homology	
					DST nucleotide range (bp #)	Database nucleotide range (bp #)
25	IMX 29	TTAC 201	Human low density lipoprotein receptor gene, exon 18 (L00352)	97	1 – 146	753 – 899
26	IMX 34	CTCC 257	Human Mitochondrial DNA control region (M76299)	99	1 – 201	450 – 650
27	IMX 35	CCAC 303	Human HepG2 3' region MboI cDNA, clone hmd3a03m3 (D17187)	99	1 – 155	88 – 243
28	IMX 36	AAGA 356	Human liver fatty acid binding protein (L-FABP) mRNA, complete cds (M10617)	99	1 – 303	89 – 391
29	IMX 37	ACTA 116	H. sapiens mitochondrial genome (X62996)	98	1 – 64	14108 – 14171
30	IMX 38	TGTG 164	H. sapiens gene TAP1, TAP 2, LMP2, LMP7 and DOB (X66401)	98	1 – 111	19074 – 19184
31	IMX 41	TCCT 85	EST H. sapiens cDNA clone IMAGE:1558956 3' similar to contains LTR5.b3 LTR5 repetitive element; mRNA sequence (AA917759)	94	1 – 32	1 – 32
6	IMX 43	GGCC 374	H. sapiens mRNA for interferon-induced 17kDa membrane protein (X84958)	99	1 – 319	329 – 647
7	IMX 45	CGAT 346	EST Jia bone marrow stroms Homo sapiens cDNA clone 1091412 3' similar to WP:CE03924 BETA TRANSDUCIN (AA599195)	99	22 – 295	8 – 281
34	IMX 46	CGGG 230	EST Homo sapiens cDNA clone IMAGE: 1117847 (AA618529)	99	1 – 178	5 – 183
35	IMX 47	CCAG 260	EST Strategene ovarian cancer (#937219) Homo sapiens cDNA clone 595329 3' (AA173988)	100	1 – 183	1 – 183
2	IMX 48	CCTT 346	NOVEL	NA	NA	NA
36	IMX 50	TCCA 371	EST qu13e05.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMA mRNA sequence (AI283282)	99	1 – 313	9 – 321
37	IMX 53	TGCC 265	H. sapiens mRNA for metallothionein isoform 2 (X97260)	99	1 – 209	120 – 327
38	IMX 54	CACT 344	Human variant urokinase plasinogen activator receptor (uPAR2) mRNA, partial cds (U09347)	100	1 – 282	10 – 291
39	IMX 58	GCGG 164	H. sapiens mRNA for leukocyte antigen CD97 (X84700)	98	1 – 113	2793 – 2905

GSP = Genome Sequencing Project, EST = Expressed Sequence Tag, NA = Not Applicable

TABLE 4: VERIFIED CANDIDATE MATCHES				
Clone ID	**Digital Address (Msp1)**	**Gene Identity (Accession #)**	**Extended Primer Sequence**	**Extended Primer Seq ID**
IMX 3	GAAT 172	Homo sapiens K12 protein precursor mRNA, complete cds. (U77643)	GATCGAATCC GGGAATCCCC CTTCAGTTCT	51
IMX 51	AGCT 341	Human interferon-induced leucine zipper protein (IFP35) mRNA, partial cds. (U72882)	GATCGAATCC GGAGCTGCAT GACGTCCTGG	52
IMX 57	GACT 96	Human mRNA for proteasome activator hPA28 subunit beta, complete cds. (D45248)	GATCGAATCC GGGACTAGAA GGAAAATAAA	53
IMX 59	GTAG 454	Human p78 protein mRNA, complete cds. (M33882)	GATCGAATCC GGGTAGCCAC TGGACTGACG	54
IMX 61	CGCT 264	Homo sapiens tumor suppressing STF cDNA 3 (TSSC3) mRNA, complete cds. (AF019953)	GATCGAATCC GGCGCTCTCC AGCCCTGCAG	55
IMX 2 (SEQ ID NO: 4)	GAAT 140	Homo sapiens K-12 protein precursor mRNA, complete cds. (U77643)	GATCGAATCC GGGAATCCCC CTTCAGTTCT	56
IMX 53 (SEQ ID NO: 37)	TGCC 265	H. sapiens mRNA for metallothionein isoform 2 (X97260)	GATCGAATCC GGTGCCAAGT GTGCCCAGGG	57

EXAMPLE 2

RT-PCR Validation

RT-PCR validation of cloned DSTs was performed, and results are presented in Table 2. RT-PCR validation of RACE clones was also performed, with results presented in Table 10. The starting amount of template was chosen based on a control curve that was generated to accurately define the linear range of amplication for the given cloned DST. Based upon the intensity of the peak cloned from the TOGA™ panel, the following amounts of template were chosen:

TOGA peak intensity	Low [cDNA]	High [cDNA]
0 – 400	2000 pg	5000 pg
300-1000	400 pg	2000 pg
> 1000	50 pg	250 pg

The PCR primers used for validation of cloned DSTs are listed in Table 7. Duplicate reaction mixtures were assembled using the appropriate low and high concentration of cDNA template chosen from the time point sample showing the strongest TOGA™ signal. The reaction mixtures were cycled for 23, 26, 29, 32, 35, and 38 cycles. The resulting amplification products from the duplicate reactions were quantitated and plotted against the cycle number to generate a standard curve, shown in tabular form below for one RACE clone, IMX 9R1. From these data, the cycle number and cDNA concentration combination which yielded acceptable levels of PCR product within the linear range of amplification were chosen for RT-PCR validation across the various time-points.

TABLE 5
IMX 9R1 Standard Curve

Cycle number	400 pg cDNA	2000 pg cDNA
23	0	119
26	94	345
29	279	535
32	652	1048
35	1219	1513
38	1421	1858

The RT-PCR validation consisted of assembling triplicate reactions using the chosen concentration of cDNA cycled to the defined cycle number. An internal control primer pair amplified under the same conditions was also performed to provide the basis for normalizing any differences between the cDNA templates. The results obtained were consistent with the initial TOGA findings, and are shown in Table 10. An example of the results from RT-PCR validation for an extended sequence obtained by the RACE method is shown in Table 6, below.

TABLE 6

	IMX 9R1	**Time**	**PCR product formed (relative intensity units)**
Conditions:	29 cycles	0	259
	400 pg cDNA	4	454
		24	1635
		44	1666

TABLE 7: RT-PCR Validation Primers

Clone ID	Digital Address (MspI)	Database Match (Accession #)	5' RT-PCR Primer and 3' RT-PCR Primer
IMX 1	GAAT 131	Homo sapiens K-12 protein precursor mRNA, complete cds. (U77643)	CCGGGAATCC CCCTTCAGTT CTT (SEQ ID NO: 66) TTTTTAGAGT CGGTTTGTTT TCTTCAT (SEQ ID NO: 67)
IMX 5	AGTC 136	Human 1-8U gene from interferon-inducible gene family (X57352)	CGGAGTCGAG TCCTGTATCA G (SEQ ID NO: 68) TCACCAGAAA CACGTGCACT TTA (SEQ ID NO: 69)
IMX 7	GCCT 97	Human mRNA for protease activator hPA28 subunit beta, complete cds (D45248/HUMPHPA28A)	TCAGAACGTG GGGATGCTGT GG (SEQ ID NO: 70) GCTCAGCATA GAAGGCCCTC AGGTC (SEQ ID NO: 71)
IMX 8	ACAG 154	EST Soares NhHMPu S1 Homo sapiens cDNA clone (AA252172)	ACAGACGGCA GCAGTCCCAG C (SEQ ID NO: 72) TTCGTGAGGA AGGGCTCATG C (SEQ ID NO: 73)
IMX 11	TGAA 134	Human LMP7 gene, exons 1-7 (L11045)	CCGGTGAAGT GCATCTTCTG C (SEQ ID NO: 74) TTTTTCTACT TAGTGGGGCA CC (SEQ ID NO: 75)
IMX 13	ACTG 191	EST Soares total fetus Nb2HF8 9w Homo Sapiens cDNA clone 784238 3' (AA446881)	GACTGCCCTC CCTCACAACA ACTGTT (SEQ ID NO: 76) CAGAGCTGGG TTCACGGTGA TGGGC (SEQ ID NO: 77)
IMX 14	AGCG 149	H. sapiens CpG island DNA genomic Mse1 fragment, clone 43f4, reverse read cpg43f4.rtla (Z65513)	CCGGAGCGAG GAGAGTAGCA CTCT (SEQ ID NO: 78) TTCTCCATGG ATGACCCCCC TCA (SEQ ID NO: 79)
IMX 16	TGAA 160	H. sapiens genes TAP1, TAP2, LMP2, LMP7 and DOB (X66401)	GGTGAAGTGC ATCTTCTGCG TGT (SEQ ID NO: 80) TTTTTCTTTA TTCTACTTAG TGGGGCAC (SEQ ID NO: 81)
IMX 17	GGGC 306	Homo sapiens mRNA for proteosome subunit MECl-1 (Y13640)	CGGACACTGA GCTCACCCACA GA (SEQ ID NO: 82) CCATAGCCTG CACAGTTTCC TCCACTA (SEQ ID NO: 83)
IMX 18	CACC 336	H. sapiens mRNA for metallothionein isoform 2 (X97260)	CTGCAAAGGG GCGTCGGACA A (SEQ ID NO: 84) CAAACGGTCA CGGTCAGGGT TGT (SEQ ID NO: 85)
IMX 19	CCCG 250	Human creatine kinase isoenzyme CK-B gene, exon 8 (M21243)	CCCTGCTGCT TCCTAACTTA TTGCCT (SEQ ID NO: 86) CTGGCAGGCC GAAACCCTAG TTTATT (SEQ ID NO: 87)
IMX 20	GTGC 257	Human mRNA for complement component C1r (X04701)	CCGGGTGCAG CAGGGGCTAT GGC (SEQ ID NO: 88) AATAGAGACT CTTAGTGGTT ATCAAGAACT GG (SEQ ID NO: 89)

TABLE 7: RT-PCR Validation Primers (Continued)

Clone ID	Digital Address (Msp1)	Database Match (Accession #)	5' RT-PCR Primer and 3' RT-PCR Primer
IMX 21A	CTTC 364	EST nn44f07.s1 NCI_CGAP_GC5 Homo sapiens cDNA clone IMAGE:1086757 similar to contains L1.t3 L1 repetitive element; mRNA sequence (AA582172)	CCTCACCCTA TGTCCTTTAT TTACA (SEQ ID NO: 90) CACAATTGCC ACAAGAAGA CTAA (SEQ ID NO: 91)
IMX 22	TTCA 420	NOVEL	CATCAAAACA CTGGAAGCCA TGGTAG (SEQ ID NO: 92) CTAGCCAGGA GGAGAGTTTC ATATGAGA (SEQ ID NO: 93)
IMX 23	ATAT 318	Homo sapiens CLN3 gene, complete CDS (X99832)	GTGAGCCGAG ATTGCAGTAC TGT (SEQ ID NO: 94) ATGTTGTCCA GTCTGGTCTC AAG (SEQ ID NO: 95)
IMX 24	AGCA 261	Human complement component C4A gene, exons 10 through 41 (M59815/HUMC4AA2)	CCCAGCTCAA CGACTTCCTC CAG (SEQ ID NO: 96) TGTTCATGAA AGCGGACACA GCAGT (SEQ ID NO: 97)
IMX 25	GCCC 379	EST Homo sapiens cDNA clone 122225 3' similar to Contains MER22 repetitive element (T98760)	ACTGTAATCA GGTGGCGCTT GGTTC (SEQ ID NO: 98) ACCTGAATTT TCAGCATACA AGGGTCAC (SEQ ID NO: 99)
IMX 26	TCCG 333	Homo sapiens transcription regulator protein (BACH 1) (AF026199)	TGCAGCAGTT ACTTCCACTC AAG (SEQ ID NO: 100) AACTGAAAGC AGGATTCCTC AAT (SEQ ID NO: 101)
IMX 29	TTAC 201	Human low density lipoprotein receptor gene, exon 18 (L00352)	GGTTACCTTG GCCGTGAGGA CAC (SEQ ID NO: 102) CTCTGAACTG AGAAATGCAA GGAGACC (SEQ ID NO: 103)
IMX 34	CTCC 257	Human Mitochondrial DNA control region (M76299)	CCAGCGTCTC GCAATGCTAT CGC (SEQ ID NO: 104) CGACATCTGG TTCCTACTTC AGGGCC (SEQ ID NO: 105)
IMX 35	CCAC 303	Human HepG2 3' region MboI cDNA, clone hmd3a03m3 (D17187)	CGGCCACCAT CTTTGTACAT GTGT (SEQ ID NO: 106) GGTCCAAGAT AATCCTATTT AGGCTACCAA (SEQ ID NO: 107)
IMX 36	AAGA 356	Human liver fatty acid binding protein (L-FABP) mRNA, complete cds (M10617)	AGAGCTCATC CAGAAGGGGA AGGA (SEQ ID NO: 108) ATGGTATTGG TGATTATGTC GCCGTT (SEQ ID NO: 109)
IMX 37	ACTA 116	H. sapiens mitochondrial genome (X62996)	CCGGACTACA ACCACGACCA (SEQ ID NO: 110) TTTTGTGTTC TTGTAGTTGA AATACA (SEQ ID NO: 111)

TABLE 7: RT-PCR Validation Primers (Continued)

Clone ID	Digital Address (Msp1)	Database Match (Accession #)	5' RT-PCR Primer and 3' RT-PCR Primer
IMX 38	TGTG 164	H. sapiens gene TAP1, TAP 2, LMP2, LMP7 and DOB (X66401)	CCGGTGTGGA CCATCGAGTC A (SEQ ID NO: 112) AAGCCCTAGA GAGTTTATTA CAAA (SEQ ID NO: 113)
IMX 45	CGAT 346	EST Jia bone marrow stroms Homo sapiens cDNA clone 1091412 3' similar to WP:CE03924 BETA TRANSDUCIN (AA599195)	GCGGCGGCGT TATTCCCATG ATG (SEQ ID NO: 114) GGTGCTCGCT TTATTTTTCA TGTGCAA (SEQ ID NO: 115)
IMX 46	CGGG 230	EST Homo sapiens cDNA clone IMAGE: 1117847 (AA618529)	CCGGCGGGAC GTCAGGGTCT ACA (SEQ ID NO: 116) GTTCCCAGGC AACCCTCCCT CGA (SEQ ID NO: 117)
IMX 47	CCAG 260	EST Strategene ovarian cancer (#937219) Homo sapiens cDNA clone 595329 3' (AA173988)	CCGGCCAGTA CGAAAAGGAC TTC (SEQ ID NO: 118) TAGATCCATA AGGCTTTAGT TCC (SEQ ID NO: 119)
IMX 48	CCTT 346	NOVEL	CCGGCCTTAA GGTTCCTCTG ACA (SEQ ID NO: 120) GGTAGGCAGC CTGAGTTGCAG TTAC (SEQ ID NO: 121)
IMX 50	TCCA 371	EST qu13e05.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMA mRNA sequence (AI283282)	CATGGTCATC ACTGGTCTGT CTG (SEQ ID NO: 122) CCTAGAGGCA TCCCAGAGAT AGCT (SEQ ID NO: 123)
IMX 54	CACT 344	Human variant urokinase plasinogen activator receptor (uPAR2) mRNA, partial cds (U09347)	CTGCCCTGAG ACCACCATGC TGT (SEQ ID NO: 124) GGGCTACATG TCCAAGGTGG CTTCTT (SEQ ID NO: 125)
IMX 58	GCGG 164	H. sapiens mRNA for leukocyte antigen CD97 (X84700)	GGCAGCTATA GTCTGGCACC AA (SEQ ID NO: 126) AGTTCTGCAG CAAAGACATG AGAG (SEQ ID NO: 127)

TABLE 8: RACE PRIMERS				
Extended Sequence Seq ID	**Clone ID**	**Database Match (Accession #)**	**1st Gene Specific Primer and Nested Race Primer**	**Primer SEQ ID**
40	IMX 6R1	EST qk33b09.x1 NCI_CGAP_Co8 H. sapiens cDNA clone IMAGE:1870745 3' similar to SW:SYS_HUMAN P49591 SERYL-TRNA SYNTHETASE (AI245918)	TTTAGGGAGA ACCTCAGAGG C	58
			AGAGGCGGGG TTCAAGC	59
41	IMX 9R1	Human interferon-induced 17-kDa/15-kDa protein mRNA, complete cds. (M13755)	TTTTTTTTGT CTTTACAACA GCC	60
			TACAACAGCC TTTATTTCCG G	61
42	IMX 12R1	EST Soares parathyroid tumor NbHPA H. sapiens cDNA clone 1360348 3', mRNA sequence (AA781258)	TTTTTTTTAG TAAACAACCT AGTC	62
			TTTTATTTGC CCAAAGCCGG	63
43	IMX 41R1	EST qn57f08.x1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1902375 3' similar to contains LTR5.b2 LTR repetitive element; mRNA sequence (AI302088)	TTTTTTTTAG CGAAAGAACA GTG	64
			AACAGTGGGC CCAGGACCGG	65

EST = Expressed Sequence Tag

TABLE 9						
Seq ID	**Clone ID**	**DST Match**	**Database Match (Accession #)**	**% Homology**	**Nucleotide homology**	
					DST nucleotide range (bp #)	**Database nucleotide range (bp #)**
40	IMX 6R1	133	EST qk33b09.x1 NCI_CGAP_Co8 H. sapiens cDNA clone IMAGE:1870745 3' similar to SW:SYS_HUMAN P49591 SERYL-TRNA SYNTHETASE (AI245918)	98	1 – 152	30 – 181
41	IMX 9R1	267	Human interferon-induced 17-kDa/15-kDa protein mRNA, complete cds. (M13755)	99	1 – 286	342 – 627
42	IMX 12R1	194	EST Soares parathyroid tumor NbHPA H. sapiens cDNA clone 1360348 3', mRNA sequence (AA781258)	99	1 – 212	24 – 235
43	IMX 41R1	363	EST qn57f08.x1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1902375 3' similar to contains LTR5.b2 LTR repetitive element; mRNA sequence (AI302088)	98	1 – 381	14 – 394

EST = Expressed Sequence Tag

TABLE 10

RACE Clone Validation

Seq ID	Clone ID	Database Match (Accession #)	Relative RACE Clone Amount*				Validation Method
			0 hour	4 hour	24 hour	44 hour	
40	IMX 6R1	EST qk33b09.x1 NCI_CGAP_Co8 H. sapiens cDNA clone IMAGE:1870745 3' similar to SW:SYS_HUMAN P49591 SERYL-TRNA SYNTHETASE (AI245918)	100	183	146	128	NBA
41	IMX 9R1	Human interferon-induced 17-kDa/15-kDa protein mRNA, complete cds. (M13755)	259	454	1635	1666	R
42	IMX 12R1	EST Soares parathyroid tumor NbHPA H. sapiens cDNA clone 1360348 3', mRNA sequence (AA781258)	617	681	562	512	R
43	IMX 41R1	EST qn57f08.x1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1902375 3' similar to contains LTR5.b2 LTR repetitive element; mRNA sequence (AI302088)	396	325	416	434	R

EST = Expressed Sequence Tag, R = Reverse Transcriptase-Polymerase Chain Reaction, NBA = Northern Blot Analysis

*Relative RACE Clone Amount = Relative RACE Clone Amount for Northerns have been corrected for background and normalized to the hybridization signal of cyclophilin. The values were obtained by exposing the Northern blot to a phosphorimaging screen and quantitated using the Phosphorimager SI.

EXAMPLE 3

Determination of Extended Sequences Using RACE

The sequences of several DSTs were extended using *Gibco BRL Life Technologies* 5' RACE (Rapid Amplification of cDNA Ends) system using primers described in Table 8. Briefly, for each DST selected, two DST-specific primers were designed. The first, a gene-specific primer, was located close to the poly-A tail. The second, overlapping the first, but more toward the 5' end of the DST was a nested RACE primer. First-strand cDNA synthesis was performed on the mRNA corresponding to the time point that displayed the highest level of expression of that individual DST, based on intensity of the fluorescently labeled PCR product. The first-strand cDNA reactions were initiated using the 3'-most DST-specific primer (closest to the poly-A tail) and SuperScript II reverse transcriptase, in the buffer and under the conditions described by the manufacturer.

After treating with RNAse, the first-strand cDNA was purified using Amicon Microcon concentrators. A homopolymeric C-tail was then added to the 3' end of the purified first-strand cDNA with dCTP and terminal deoxynucleotidyl transferase. Two serial PCR reactions were then performed for each DST. In the first PCR, the C-tailed first-strand cDNA was used as the template with the same 3'-most DST-specific primer as used in the beginning step, in conjunction with and an anchor primer ("AP" - provided in the kit by the manufacturer). After PCR amplification, the resultant PCR reaction products were diluted and a small portion of this used as template in a second PCR reaction. The primers for this second PCR reaction were a second DST-specific primer (located more towards the 5' end of the DST) and a second primer provided in the kit ("AUAP" - abridged upstream anchor primer).

PCR products from both of the serial PCR reactions were analyzed on acrylamide gels. An aliquot of each of the products was then ligated into *Invitrogen's* TOPO II vector and transformed into TOP 10 cells. Bacterial colonies were picked and grown overnight for mini-preparation of the plasmid DNA. The plasmid DNAs were then digested with a restriction enzyme to determine the presence and size of a PCR insert. The insert in the plasmid DNA was then sequenced and analyzed. The RACE derived sequences for a specific DST were aligned both with each other and with the original DST sequence. Due

to the design of the DST-specific PCR primers, RACE products generated from the second serial PCR reaction only contained a portion of the original DST sequence. Validation of the RACE-PCR generated clones was performed by RT-PCR (see Table 10) as described above in Example 2. The database matches and alignments to DSTs are presented in Table 9, above. The results of validation by Northern blot analysis (IMX 6R1) or RT-PCR (IMX 9R1, IMX 12R1 and IMX 41R1) are presented in Table 10, above.

EXAMPLE 4

Further Studies of IMX1

As noted above in Tables 2 and 3, IMX 1 was found to have a region of high homology to human K12 mRNA. Further studies are described below, and in Lyman, S.D., et al., Identification of CD7 as a cognate of the human K12 (SECTM1) protein, *J. Biol. Chem.* 275(2):3431-3437 (2000), which is incorporated by reference herein.

Regulation of the Murine K12 Gene in Mouse Models of IBD

Changes in K12 expression were measured in various mouse models of inflammatory bowel disease (IBD). There was a modest downregulation of K12 mRNA in DSS-induced colitis in C57Bl/6 mice and in anti-CD3 antibody-induced ileitis in Balb/c mice. There was also a downregulation of K12 mRNA in mdrla transgenic knockout animals that had developed IBD, compared to knock-out animals that were not sick, the latter having less K12 mRNA than wild-type parental mice. K12 was administered to mice undergoing anti-CD3 induced ileitis. No effect was seen.

Effect of K12 Protein on T84 Cell Barrier Formation and Breakdown

In tests were performed to determine the effect of K12 in the T84 barrier function model, no effect was seen on either the initial formation of the barrier, or its subsequent breakdown in response to IFN-gamma treatment.

Effect of K12 Protein on Murine Endotoxemic Mouse Model

The murine K12-Fc fusion protein was tested in an LPS/D-gal septic shock model. A recent paper from showed that CD7 knockout mice are resistant to endotoxic shock in the LPS/D-gal model (Sempowski et al. J. Exp. Med. 189: 1011-1016 (1999). Since K12 is a ligand for CD7 is was possible that the K12 protein might either exacerbate or ameliorate the endotoxic effects of LPS. In the first experiment, K12-Fc administration (50 μg/mouse) did not appear to enhance the toxicity of 5 ng LPS in combination with D-gal, producing about 30% mortality in both groups. In a second experiment, 50 μg/mouse of K12-Fc appeared to offer some protection against LPS challenge in this model, with 50% survival as compared to 20% in the control group (N=10 mice/group). However, in a third and larger experiment (N=15/group) K12-Fc (this time at 100 μg/mouse) failed to block the toxicity of LPS when given in combination with 30 ng LPS + D-gal (about 40% survival in both groups). These experiments failed to support a role for the Kl2 protein in LPS mediated shock in mice. However, several differences in the experimental protocol from that used by Sempowski et al., a different strain of mice, lower LPS dose and different route of injection prevent a simple comparison.

Effect of K12 Protein on T Cell Proliferation and Cytokine Secretion

K12 protein was tested in a variety of T cell proliferation assays, including anti-CD3 induced proliferation of naive OVA TcR transgenic lymph node cells *in vitro;* antigen induced proliferation of naive OVA TcR transgenic lymph node cells *in vitro;* antigen - induced proliferation of memory OVA TcR transgenic lymph node cells *in vitro.* No effect was seen, with one exception. K12 inhibited the concanavalin-A induced, but not the anti-T cell receptor induced, proliferation of mouse lymph node T cells. Numerous other molecules also have this effect, and its significance is unclear.

K12 was found to have no effect on human NK cell and mouse thymocyte proliferation. K12 induced CD25, CD54 and CD69 on human NK cells. Human K12 binds to the human T cell molecule CD7.

Identification of a Cognate for the K12 Protein

Mouse K12 also binds to mouse CD7. In both the human and mouse genomes, the K12 gene maps very near to the CD7 gene. See Lyman, et al. (2000) for further details and discussion.

5

We claim:

1. An isolated nucleic acid molecule comprising a polynucleotide chosen from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43.

2. An isolated polypeptide encoded by a polynucleotide chosen from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43.

3. An isolated nucleic acid molecule comprising a polynucleotide at least 95% identical to the isolated nucleic acid molecule of claim 1.

4. An isolated nucleic acid molecule at least ten bases in length that is hybridizable to the isolated nucleic acid molecule of claim 1 under stringent conditions.

5. An isolated nucleic acid molecule encoding the polypeptide of claim 2.

6. An isolated nucleic acid molecule encoding a fragment of the polypeptide of claim 2.

7. An isolated nucleic acid molecule encoding a polypeptide epitope of the polypeptide of claim 2.

8. The polypeptide of claim 2 wherein the polypeptide has biological activity.

9. An isolated nucleic acid encoding a species homologue of the polypeptide of claim 2.

10. The isolated nucleic acid molecule of claim 1, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

11. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

12. A recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

13. A method of making the recombinant host cell of claim 12.

14. The recombinant host cell of claim 12 comprising vector sequences.

15. The isolated polypeptide of claim 2, wherein the isolated polypeptide comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

16. An isolated antibody that binds specifically to the isolated polypeptide of claim 2.

17. The isolated antibody of claim 16 wherein the antibody is a monoclonal antibody.

18. The isolated antibody of claim 16 wherein the antibody is a polyclonal antibody.

19. A recombinant host cell that expresses the isolated polypeptide of claim 2.

20. An isolated polypeptide produced by the steps of:

(a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and

(b) isolating the polypeptide.

21. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 2 or the polynucleotide of claim 1.

22. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

23. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of claim 2 in a biological sample; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

24. A method for identifying a binding partner to the polypeptide of claim 2 comprising:

(a) contacting the polypeptide of claim 2 with a binding partner; and

(b) determining whether the binding partner effects an activity of the polypeptide.

25. The gene corresponding to the cDNA sequence of the isolated nuclei acid of claim 1.

26. A method of identifying an activity of an expressed polypeptide in a biological assay, wherein the method comprises:

(a) expressing the polypeptide of claim 2 in a cell;

(b) isolating the expressed polypeptide;

(c) testing the expressed polypeptide for an activity in a biological assay; and

(d) identifying the activity of the expressed polypeptide based on the test results.

27. A substantially pure isolated DNA molecule suitable for use as a probe for genes regulated in gastrointestinal inflammation, chosen from the group consisting of the DNA molecules identified in Table 1, having a 5' partial nucleotide sequence and length as described by their digital address, and having a characteristic regulation pattern in T84 cells treated with interferon-gamma.

28. A kit suitable for detecting the presence of the polypeptide of the claim 2 in a mammalian tissue sample comprising a first antibody which immunoreacts with a mammalian protein encoded by a gene corresponding to the polynucleotide of claim 1 or with a polypeptide of claim 2 in an amount sufficient for at least one assay, instructions for use and suitable packaging material.

29. A kit of claim 28 further comprising a second antibody that binds to the first antibody.

30. The kit of claim 29 wherein the second antibody is labeled.

31. The kit of claim 30 wherein the label comprises enzymes, radioisotopes, fluorescent compounds, colloidal metals, chemiluminescent compounds, phosphorescent compounds, or bioluminescent compounds.

32. A kit for suitable for detecting the presence of a gene regulated in gastrointestinal inflammation, comprising:

at least one polynucleotide of claims 1 or 4, or fragment thereof having at least 10 contiguous bases, in an amount sufficient for at least one assay;

label means;

instructions for use; and

suitable packaging material.

A

B

C

D

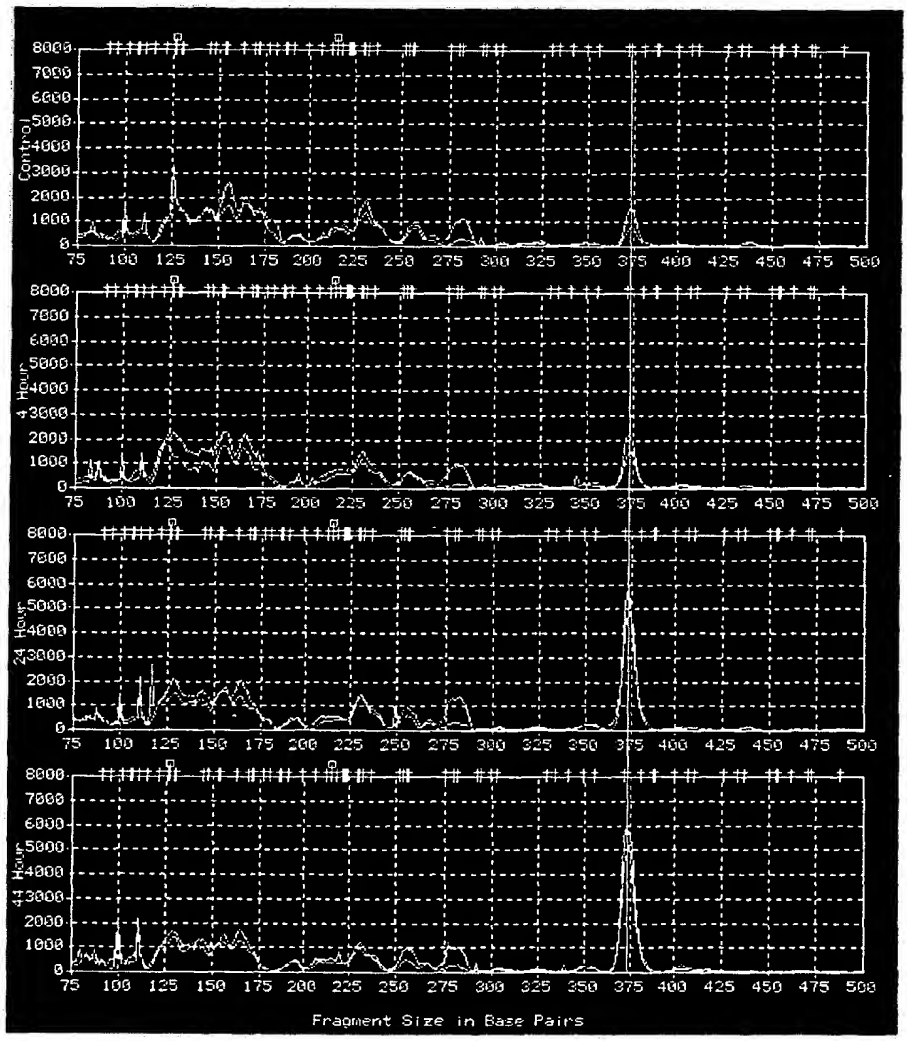
Control
4 Hours
4 Hours
4 Hours
8000
7000
6000
5000
4000
3000
2000
1000
0
75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500
Fragment Size in Base Pairs

Fig. 1

A

B

C

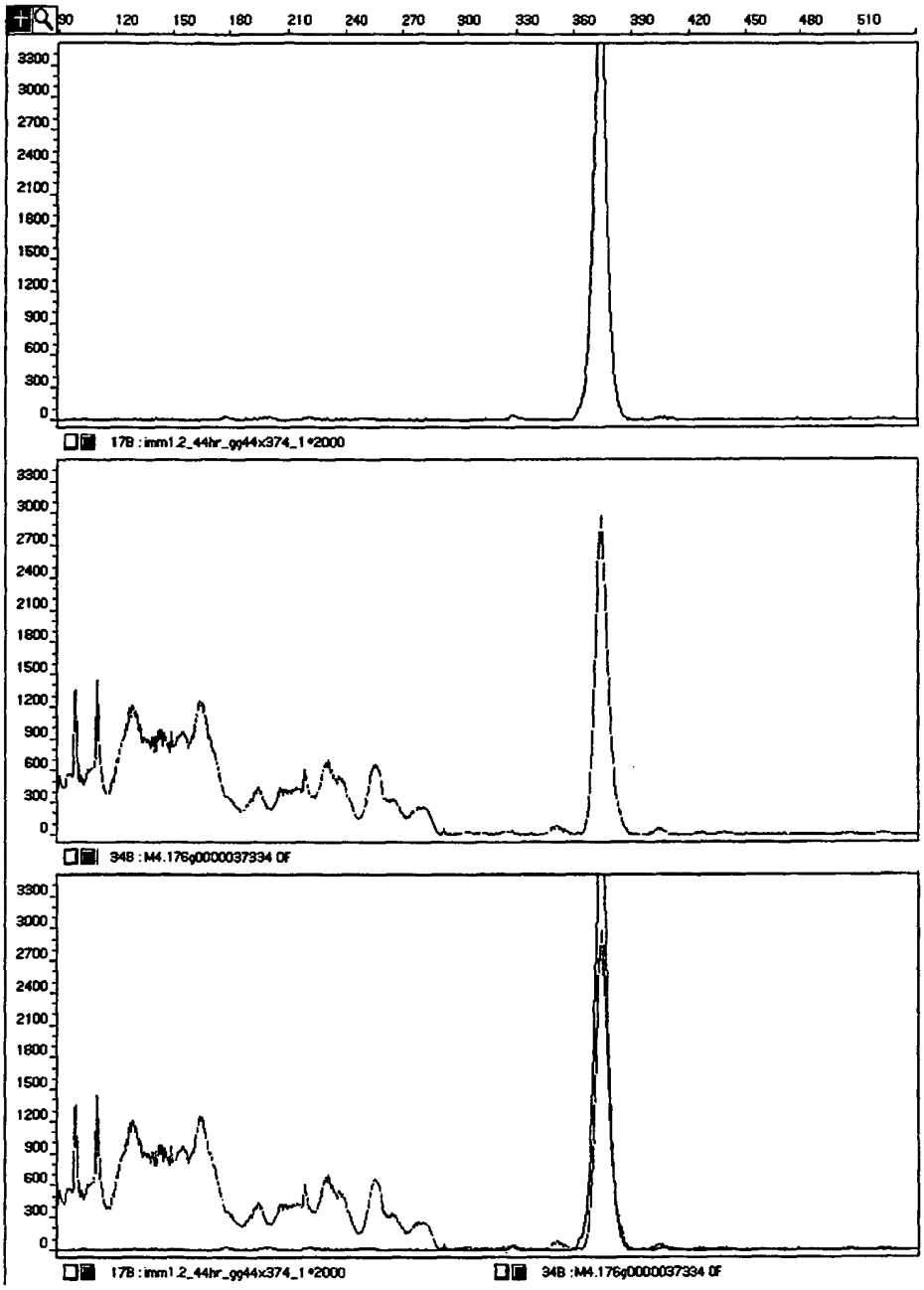
90 120 150 180 210 240 270 300 330 360 390 420 450 480 510
3300 3000 2700 2400 2100 1800 1500 1200 900 600 300 0
178 : imm1.2_44hr_gg44x374_1 *2000
348 : M4.176g0000037334 0F
178 : imm1.2_44hr_gg44x374_1 *2000
348 : M4.176g0000037334 0F

Fig. 2

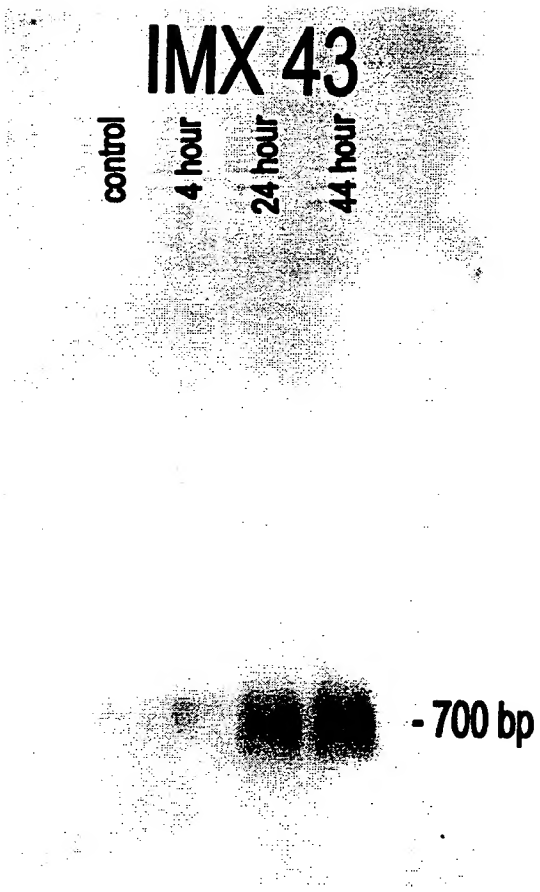

Fig. 3

SEQUENCE LISTING

<110> Youakim, Adel
DuBose, Robert
Sims, John
Pribyl, Thomas M
Hilbush, Brian S
Hasel, Karl W

<120> Modulation of Gene Expression in Gastrointestinal Inflammation

<130> 98,501-A

<140>
<141> 2000-06-01

<150> 60/137,058
<151> 1999-06-01

<160> 127

<170> PatentIn Ver. 2.0

<210> 1
<211> 370
<212> DNA
<213> Homo sapiens

<400> 1
cggttcagcc acagcaaatg tatagagtgc tacaaccctc tatacaggga tacagcaatg 60
tgaaaccaga ggtcatgcat ctgaggtttc atcaaaacac tggaagccat ggtagttatt 120
gtccttgata ggtatgtcaa cctcaacacc aatgtccttc ctggctgatg gagaaatggg 180
ctgcccgatt cacagacttc ttgatctttc ctcagcatat gatactctca tatgaaactc 240
tcctcctggc taggagtggc agcctacatc gctaatccca gcaccttggg aagctgaggc 300
aggaggatta ctggaggcca ggagtttgag accaggctgg gcaatacagc gagactctct 360
ctcctaaaaa 370

<210> 2
<211> 294
<212> DNA
<213> Homo sapiens

<400> 2
cggccttaag gttcctctga cagttgtctc aagaagatca ggatttgagc atctcttact 60
tcaaattatt cttccagtgg ttttgttatc ttctcaatgt tatagtgaaa tataataata 120
catgcagaaa aaagcataaa tcaactattg aataacgaaa atgtgaacct ctatgtaact 180
gcaactcagg ctgcctaccg tttttaattt ctctctaatc tttaaaaaca aatactatct 240
ttactaaaga tgtttcagag taagataaaa tacaaagaga gtaattccca aaaa 294

<210> 3
<211> 79
<212> DNA
<213> Homo sapiens

<400> 3
cgggaatccc ccttcagttc tttgaaaaag ttccatgact cgaatatctg aaatgaagaa 60
aacaaaccga ctctaaaaa 79

<210> 4
<211> 89
<212> DNA
<213> Homo sapiens

<400> 4
cgggaatccc ccttcagttc tttgaaaaag ttccatgact cgaatatctg aaatgaagaa 60
aacaaaccaa aaaaaaaaaa aaagaaaaa 89

<210> 5
<211> 88
<212> DNA
<213> Homo sapiens

<400> 5
cggagtcgag tcctgtatca gcccttttatc ctcacacgct tttctacaat ggcattcaat 60
aaagtgcacg tgtttcttgg tacaaaaa 88

<210> 6
<211> 323
<212> DNA
<213> Homo sapiens

<400> 6
cggggcccag gcctatgcct ccaccgccaa gtgcctgaac atctgggccc tgattctggg 60
catcctcatg accattggat tcaccctgtt actggtattc ggctctgtga cagtctacca 120
tattatgtta cagataatac aggaaaaacg gggttactag tagccgccca tagcctgcaa 180
cctttgcact ccactgtgca atgctggccc tgcacgctgg ggctgttgcc cctgcccccc 240
tggtcctgcc cctagataca gcagtttata cccacacacc tgtctacagt gtcattcaat 300
aaagtgcacg tgcttgtgaa aaa 323

<210> 7
<211> 294
<212> DNA
<213> Homo sapiens

<400> 7
cggcgatgcg gcggcgttat tcccatgatg cctctgtcaa ggagtggaca atcacctact 60
gaggagcccc acccccgcct ctggatggac cgaatcaggg actagagttt aactgcagcg 120
gaacatgtca tttctctatt tctgtgacgc gccccccatgc ccccacccca ccacaagagg 180
caggagggcc cagtcatgac cctcgtctct gcagggtgtc tgtacacgtt cttctgaaag 240
ctttagacag taacagtttg cacatgaaaa ataaagcgag cacctagaca aaaa 294

<210> 8
<211> 103
<212> DNA
<213> Homo sapiens

<400> 8
cggacagacg gcagcagtcc cagctctggt ttccttctcg gtttattctg ttagaatgaa 60
atggttccca taaataaggg gcatgagccc ttcctcacaa aaa 103

<210> 9
<211> 30
<212> DNA
<213> Homo sapiens

<400> 9
cggaaataaa ggctgttgta aagacaaaaa 30

<210> 10
<211> 85
<212> DNA
<213> Homo sapiens

<400> 10
cggtgaagtg catcttctgc gtgttctcta tttgaacaag catttccccc agggaagttt 60
ctgggtgccc cactaagtag aaaaa 85

<210> 11
<211> 40
<212> DNA
<213> Homo sapiens

<400> 11
cggctttggg caaataaaag actaggttgt ttactaaaaa 40

<210> 12
<211> 138
<212> DNA
<213> Homo sapiens

<400> 12
cggactgccc tccctcacaa caactgttgt tcaggctaca ccaaagagtc ctccgttaaa 60
acccattcaa ccaaagccca caattatggg agagcccatc accgtgaacc cagctctggt 120
gtcactcaaa aacaaaaa 138

<210> 13
<211> 98
<212> DNA
<213> Homo sapiens

<400> 13
cggagcgagg agagtagcac tcttgtgcgg gatattgatt tcacggagga tggtggtcaa 60
gggaccccta tctgaggggg gtcatccatg gagaaaaa 98

<210> 14
<211> 108
<212> DNA
<213> Homo sapiens

<400> 14
cggtgaagtg catcttctgc gtgttctcta tttgaacaag catttcccac agggaagttt 60
ctgggtgccc cactaagtag aataaagaaa aacggttata agtaaaaa 108

<210> 15
<211> 255
<212> DNA
<213> Homo sapiens

<400> 15
cgggggcaat gtggacgcat gtgtgatcac aaagactggc gccaagctgc tgcggacact 60
gagctcaccc acagagcccg tgaagaggtc tggccgctac cactttgtgc ctggaaccac 120
agctgtcctg acccagacag tgaagccact aaccctggag ctagtggagg aaactgtgca 180
ggctatggag gtggagtaag ctgaggctta gagcttggaa caaggggaa taaacccaga 240
aaatacagtc aaaaa 255

<210> 16
<211> 285

<212> DNA
<213> Homo sapiens

<400> 16
cggcacctgc aaatgcaaag agtgcaaatg cacctcctgc aagaaaagct gctgctcctg 60
ctgccctgtg ggctgtgcca agtgtgccca gggctgcatc tgcaaagggg cgtcggacaa 120
gtgcagctgc tgcgcctgat gctgggtcag ccccgctccc agatgtaaag aacgcgactt 180
ccacaaacct ggatttttta tgtacaaccc tgaccgtgac cgtttgctat attccttttt 240
ctatgaaata atgtgaatga taataaaaca gctttgactt aaaaa 285

<210> 17
<211> 200
<212> DNA
<213> Homo sapiens

<400> 17
cggcccgcac ccgacaccag ccctgctgct tcctaactta ttgcctgggc agtgcccacc 60
atgcacccct gatgttcgcc gtctggcgag cccttagcct tgctgtagag acttccgtca 120
cccttggtag agtttatttt tttgatggct aagatactgc tgatgctgaa ataaactagg 180
gtttcggcct gccagaaaaa 200

<210> 18
<211> 203
<212> DNA
<213> Homo sapiens

<400> 18
cgggtgcagc aggggctatg gcttctacac caaagtgctc aactacgtgg actggatcaa 60
gaaagagatg gaggaggagg actgagccca gaattcacta ggttcgaatc cagagagcag 120
tgtggaaaaa aaaaacaaaa aacaactgac cagttgttga taaccactaa gagtctctat 180
taaaattact gatgcaggaa aaa 203

<210> 19
<211> 316
<212> DNA
<213> Homo sapiens

<400> 19
cggcttcctc accctatgtc ctttatttac aaagctgtgc tcctattcat gagcatggaa 60
tgtttttcca tttgtttgtg acatctctta tttctttcag gggtatcttg taattctcat 120
tatatatatc ttttgcttcc ttggttagct gtatttttag gtatttttagt cttcttgtgg 180
caattgtgaa tgggattgca ttcctgattt ggctcttggc ttgaatgtta ttaacgccac 240
atttttaaa tagacaaaaa tatgagatta aaaatgttga attttaccaa caataaaagt 300
tgttcaaagg gaaaaa 316

<210> 20
<211> 265
<212> DNA
<213> Homo sapiens

<400> 20
cggatatgga ggttgcagtg agccgagatt gcagtactgt actccagcct gggccacaga 60
gcgagactcc atctccaaaa aaaagactta ggaactcaat tctgtgatag caaaatacag 120
aaatagagat tatactttag gctaggtgtt gtggctcaca cctgtaatcc caatgctttt 180
ggaggctagg gtgggaggat cactcgaagc caggagcttg agaccagact ggacaacata 240
gtgagacccc ccacctccac aaaaa 265

<210> 21
<211> 209
<212> DNA

<213> Homo sapiens

<400> 21
cggagcaccc gccagcgggc agcctgtgcc cagctcaacg acttcctcca ggagtatggc 60
actcaggggt gccaggtgtg agggctgccc tcccacctcc gctgggagga acctgaacct 120
gggaaccatg aagctggaag cactgctgtg tccgctttca tgaacacagc ctgggaccag 180
ggcatattaa aggcttttgg catcaaaaa 209

<210> 22
<211> 321
<212> DNA
<213> Homo sapiens

<400> 22
cgggcccact gtaatcaggt ggcgctggtt ctatctcagg ttggggacct tagcttttct 60
aggctgaaag aatggagccc ttctgttcag tggtgtccat ctgggccctg gactctggat 120
ttgaccgagg ccctggaggg gagggccatg gagttgtgct tgtgtgtcat gtgcacggtc 180
ctggtttact gtgcaccttc tctaactaga tccttagcca agggcttcac atacagcgtg 240
gttatgttta ttaatgagtc tgtcttatga agtgaccctt gtatgctgaa aattcaggta 300
tatttgtacc aaagataaaa a 321

<210> 23
<211> 283
<212> DNA
<213> Homo sapiens

<400> 23
cggtccgtgc tggcggcatg cagcagttac ttccactcaa gaatcgtagg ccaggctgat 60
ggagagctga acattactct tccagaagag gtgacagtta aaggatttga acctttaatt 120
cagtttgcct acactgctaa actgatttta agtaaagaga atgtggatga agtgtgcaaa 180
tgtgtggagt ttttaagtgt acataatatt gaggaatcct gctttcagtt tctgaaattt 240
aagtttttgg actccactgc agaccagcaa gaatgcccaa aaa 283

<210> 24
<211> 300
<212> DNA
<213> Homo sapiens

<400> 24
cggccgtgga ccatagtcgg gatgtcagca gacacacatg ggcagcccaa agctgcaggt 60
gcccagggcc caccccagcc tcgcctgtca cccccactcc cgcctcaggg ccaggcccag 120
gcctcaccac ctgacgctgc atgagacatt gacaccagaa agccctcttg tctgagggac 180
atatcagtga cctcgacgtt gcctttagac tacagttgtg ttagcctctt gcgtattggc 240
tttttcagag tcatttatga gcagaaaaaa aaaaaagtaa aactttgcta atatcaaaaa 300

<210> 25
<211> 147
<212> DNA
<213> Homo sapiens

<400> 25
cggttacctt ggccgtgagg acacgtggcc tgcacccagg tgtggctgtc aggacaccag 60
cctggtgccc atcctcccga cccctaccca cttccattcc cgtggtctcc ttgcatttct 120
cagttcagag ttgtacactg tgaaaaa 147

<210> 26
<211> 205
<212> DNA
<213> Homo sapiens

<400> 26
cggctccagc gtctcgcaat gctatcgcgt gcataccccc cagacgaaaa taccaaatgc 60
atggagagct cccgtgagtg gttaataggg tgatagacct gtgatccatc gtgatgtctt 120
atttaagggg aacgtgtggg ctatttaggc tttatggccc tgaagtagga accagatgtc 180
ggatacagtt cactttagct aaaaa 205

<210> 27
<211> 254
<212> DNA
<213> Homo sapiens

<400> 27
cggccaccat ctttgtacat gtgtctttgc agcttagtgc atctctttga caaagatttg 60
tgcaatgatg tattttattc aacacatcat tctgaaagaa cgtgtggaaa actaatgact 120
gagcacaaac ctgttgaaga tgcagttctt cttggtagcc taaataggat tatcttggac 180
ctaagatagt gaccttgttt ttttttttctt aatttttcca ttgaaataca tggccttaag 240
gaaatatgca aaaa 254

<210> 28
<211> 304
<212> DNA
<213> Homo sapiens

<400> 28
cggaagagct catccagaag gggaaggata tcaaggggt gtcggaaatc gtgcagaatg 60
ggaagcactt caagttcacc atcaccgctg ggtccaaagt ggtccaaaac gaattcacgg 120
tgggggagga atgtgagctg gagacaatga caggggagaa agtcaagaca gcggttcagt 180
tggaaggtga caataaactg gtgacaactt tcaaaaacat caagtctgtg accgaactca 240
acggcgacat aatcaccaat accatgacat tgggtgacat tgtcttcaag agaatcagca 300
aaaa 304

<210> 29
<211> 66
<212> DNA
<213> Homo sapiens

<400> 29
cggactacaa ccacgaccaa tgatatgaaa aaccatcgtt gtatttcaac tacaagaaca 60
caaaaa 66

<210> 30
<211> 113
<212> DNA
<213> Homo sapiens

<400> 30
cggtgtggac catcgagtca tcttgggcaa tgaactgcca aaattctatg atgagtgaac 60
cttccccaga cttctctttc ttattttgta ataaactctc tagggcttaa aaa 113

<210> 31
<211> 32
<212> DNA
<213> Homo sapiens

<400> 31
cggtcctggg cccactgttc tttcgctaaa aa 32

<210> 32
<211> 37

<212> DNA
<213> Homo sapiens
<400> 32
cgggcttgaa ccccgcctct gaggttctcc ctaaaaa 37

<210> 33
<211> 43
<212> DNA
<213> Homo sapiens
<400> 33
cgggcctaga aggaaaataa atgatctata tgttgtgcaa aaa 43

<210> 34
<211> 178
<212> DNA
<213> Homo sapiens

<400> 34
cggcgggacg tcagggtcta catcagcctc ctgcccctgg gcgatggact caccttggcc 60
ttcaagatct agggctggcc cctagtgagt gggctcgagg gagggttgcc tgggaacccc 120
aggaattgac cctgagtttt aaattcgaaa ataaagtggg gctgggacac accaaaaa 178

<210> 35
<211> 206
<212> DNA
<213> Homo sapiens

<400> 35
cggccagtac gaaaaggact tcctggaggc ggatgcttac ggagtttctt acaacatgaa 60
agctgacata gaagatgagc tagataaaac cattgaattg ttttctaccg aaaagaaaga 120
taaaaatgat tccgagagag aaccttccaa gaaaatagaa aaggaactaa agccttatgg 180
atctagtgcc atcaacatcc taaaaa 206

<210> 36
<211> 319
<212> DNA
<213> Homo sapiens

<400> 36
cggtccaggg aagaggctcg ctcccgccca tggtcatcac tggtctgtct gctctgttgt 60
ctgttctttc cctgactccc tcccaccgaa ggcctgatgg ctactcaccc ctctgggatg 120
gctatgggag aggaggagtg atggggaccg ccacctttc tgcaggaaat gtgcccagca 180
gctcttggtc aaagcactgt tgctataagc tatctctggg atgcctctag gccccttcc 240
ctctacacac ccctgggaaa agattacact gtattaactc tcgaggagtt tcctcaccaa 300
taaacagaca accgcaaaa 319

<210> 37
<211> 214
<212> DNA
<213> Homo sapiens

<400> 37
cggtgccaag tgtgcccagg gctgcatctg caaaggggcg tcggacaagt gcagctgctg 60
cgcctgatgc tgggacagcc ccgctcccag atgtaaagaa cgcgacttcc acaaacctgg 120
atttttatg tacaaccctg accgtgaccg tttgctatat tccttttct atgaaataat 180
gtgaatgata ataaaacagc tttgactaga aaaa 214

<210> 38

<211> 292
<212> DNA
<213> Homo sapiens

<400> 38
cggcactcac gaacgctcac tctggggaag ctggttgcca tgtaaaagta ctactgccct 60
gagaccacca tgctgtgagg aagcccaagc tactcatgta taaatgccat gtggagatag 120
agccccagat gtttcagcca tctcagccca ggcaccagac aagtgggtga agaagccacc 180
ttggacatgt agccccagca gatgtgatat agagaagaaa caggaaactt ggctatatta 240
gtttcctagg gctgcctgtg ataaattatt acaaacttta taaactcaaa aa 292

<210> 39
<211> 113
<212> DNA
<213> Homo sapiens

<400> 39
cgggcgggag gttctcactg ttgtgaaggt tgtagacgtt gtgtaatgtg tttttatctg 60
ttaaaatttt tcagtgttga cacttaaaat taaacacatg catacaggaa aaa 113

210> 40
<211> 152
<212> DNA
<213> Homo sapiens

<400> 40
aacagtgcca tgtaccacac agatgttcct gtctcctcgc atgggcatag ggacccatca 60
ttgatgactg atgaaaccat gtaataaagc atctctgggg agggcttagg actcttcctc 120
agtcttcttc cccgggcttg aaccccgcct ct 152

<210> 41
<211> 286
<212> DNA
<213> Homo sapiens

<400> 41
caagggccgc agcagcacct acgaggtgcg gctgacgcag accgtggccc acctgaagca 60
gcaagtgagc gggctggagg gtgtgcagga cgacctgttc tggctgacct tcgaggggaa 120
gcccctggag gaccagctcc cgctggggga gtacggcctc aagcccctga gcaccgtgtt 180
catgaatctg cgcctgcggg gaggcggcac agagcctggc gggcggagct aagggcctcc 240
accagcatcc gagcaggatc aagggccgga aataaaggct gttgta 286

<210> 42
<211> 212
<212> DNA
<213> Homo sapiens

<400> 42
cagtttcctc ctccacaact gaatatagtg gctgaaaact ggggagatac ttgatggcgc 60
gaatgtccgt tttctctccc ttcccacctc ctgcaggaag caggacgggg caggcagcac 120
ctggtaggca cagtgctttg cccctcctcc ccttcccttc tggaagtctt ggggcctcag 180
tgcttgcaac agccggcttt gggcaaataa aa 212

<210> 43
<211> 381
<212> DNA
<213> Homo sapiens

<400> 43
taatggaatg tctcggtgta aagctgacca ttcccattcg ttctattctg agataggaga 60
aaaccgccct gtggctggag gtgagatatg ctggcagcaa tactgctctg ttactccttg 120
ctacactgag atgtttgggt aaagagaaac ataaatctag cctacgtgca catctgggca 180
cagtaccttt ccttgaactt attcgtgata cagattcctt tgctcacatg tttccctgct 240
gaccttcttc ccacctgttg ccctgctaca ctccccctcgc taagacagta aaaataatga 300
tcaataaata ctgagggaac tcagaggcca gcgccggtgc gggtcctcca catgctgagc 360
gccggtcctg ggcccactgt t 381

<210> 44
<211> 49
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence:synthetic primer

<400> 44
gaattcaact ggaagcggcc cgcaggaatt ttttttttt ttttttvnn 49

<210> 45
<211> 16
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer

<400> 45
aggtcgacgg tatcgg 16

<210> 46
<211> 16
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence:synthetic primer

<400> 46
ggtcgacggt atcggn 16

<210> 47
<211> 15
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence:synthetic primer

<400> 47
gagctccacc gcggt 15

<210> 48
<211> 16
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence:synthetic primer

<400> 48
cgacggtatc ggnnnn 16

<210> 49
<211> 16
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer

<400> 49
cgacggtatc ggggcc 16

<210> 50
<211> 30
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer

<400> 50
gatcgaatcc ggggcccagg cctatgcctc 30

<210> 51
<211> 30
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 3 extended primer

<400> 51
gatcgaatcc gggaatcccc cttcagttct 30

<210> 52
<211> 30
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 51 extended primer

<400> 52
gatcgaatcc ggagctgcat gacgtcctgg 30

<210> 53
<211> 30
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 57 extended primer

<400> 53

gatcgaatcc gggactagaa ggaaaataaa 30

<210> 54
<211> 30
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 59 extended primer

<400> 54
gatcgaatcc gggtagccac tggactgacg 30

<210> 55
<211> 30
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 61 extended primer

<400> 55
gatcgaatcc ggcgctctcc agccctgcag 30

<210> 56
<211> 30
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 2 extended primer

<400> 56
gatcgaatcc gggaatcccc cttcagttct 30

<210> 57
<211> 30
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 53 extended primer

<400> 57
gatcgaatcc ggtgccaagt gtgcccaggg 30

<210> 58
<211> 21
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 6R1 gene specific primer

<400> 58
tttagggaga acctcagagg c 21

<210> 59
<211> 17
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 6R1 nested RACE primer

<400> 59
agaggcgggg ttcaagc 17

<210> 60
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 9R1 gene specific primer

<400> 60
tttttttttgt ctttacaaca gcc 23

<210> 61
<211> 21
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 9R1 nested RACE primer

<400> 61
tacaacagcc tttatttccg g 21

<210> 62
<211> 24
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 12R1 gene specific primer

<400> 62
tttttttttag taaacaacct agtc 24

<210> 63
<211> 20
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 12R1 nested RACE primer

<400> 63

ttttatttgc ccaaagccgg 20

<210> 64
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 41R1 gene specific primer

<400> 64
ttttttttag cgaaagaaca gtg 23

<210> 65
<211> 20
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 41R1 nested RACE primer

<400> 65
aacagtgggc ccaggaccgg 20

<210> 66
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 1 5' RT-PCR primer

<400> 66
ccgggaatcc cccttcagtt ctt 23

<210> 67
<211> 27
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 1 3' RT-PCR primer

<400> 67
tttttagagt cggtttgttt tcttcat 27

<210> 68
<211> 21
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 5 5' RT-PCR primer

<400> 68
cggagtcgag tcctgtatca g 21

<210> 69
<211> 23

<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 5 3' RT-PCR primer

<400> 69
tcaccagaaa cacgtgcact tta 23

<210> 70
<211> 22
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 7 5' RT-PCR primer

<400> 70
tcagaacgtg gggatgctgt gg 22

<210> 71
<211> 25
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 7 3' RT-PCR primer

<400> 71
gctcagcata gaaggccctc aggtc 25

<210> 72
<211> 21
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 8 5' RT-PCR primer

<400> 72
acagacggca gcagtcccag c 21

<210> 73
<211> 21
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 8 3' RT-PCR primer

<400> 73
ttcgtgagga agggctcatg c 21

<210> 74
<211> 21
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic

primer: IMX 11 5' RT-PCR primer

<400> 74
ccggtgaagt gcatcttctg c 21

<210> 75
<211> 22
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 11 3' RT-PCR primer

<400> 75
tttttctact tagtggggca cc 22

<210> 76
<211> 26
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 13 5' RT-PCR primer

<400> 76
gactgccctc cctcacaaca actgtt 26

<210> 77
<211> 25
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 13 3' RT-PCR primer

<400> 77
cagagctggg ttcacggtga tgggc 25

<210> 78
<211> 24
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 14 5' RT-PCR primer

<400> 78
ccggagcgag gagagtagca ctct 24

<210> 79
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 14 3' RT-PCR primer

<400> 79
ttctccatgg atgacccccc tca 23

<210> 80
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 16 5' RT-PCR primer

<400> 80
ggtgaagtgc atcttctgcg tgt 23

<210> 81
<211> 28
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 16 3' RT-PCR primer

<400> 81
tttttcttta ttctacttag tggggcac 28

<210> 82
<211> 22
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 17 5' RT-PCR primer

<400> 82
cggacactga gctcacccaca ga 22

<210> 83
<211> 27
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 17 3' RT-PCR primer

<400> 83
ccatagcctg cacagtttcc tccacta 27

<210> 84
<211> 21
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 18 5' RT-PCR primer

<400> 84

ctgcaaaggg gcgtcggaca a 21

<210> 85
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 18 3' RT-PCR primer

<400> 85
caaacggtca cggtcagggt tgt 23

<210> 86
<211> 26
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 19 5' RT-PCR primer

<400> 86
ccctgctgct tcctaactta ttgcct 26

<210> 87
<211> 26
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 19 3' RT-PCR primer

<400> 87
ctggcaggcc gaaaccctag tttatt 26

<210> 88
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 20 5' RT-PCR primer

<400> 88
ccgggtgcag caggggctat ggc 23

<210> 89
<211> 32
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 20 3' RT-PCR primer

<400> 89
aatagagact cttagtggtt atcaagaact gg 32

<210> 90
<211> 25

<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 21A 5' RT-PCR primer

<400> 90

cctcacccta tgtcctttat ttaca 25

<210> 91
<211> 24
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 21A 3' RT-PCR primer

<400> 91
cacaattgcc acaagaaga ctaa 24

<210> 92
<211> 26
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 22 5' RT-PCR primer

<400> 92
catcaaaaca ctggaagcca tggtag 26

<210> 93
<211> 28
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 22 3' RT-PCR primer

<400> 93
ctagccagga ggagagtttc atatgaga 28

<210> 94
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 23 5' RT-PCR primer

<400> 94
gtgagccgag attgcagtac tgt 23

<210> 95
<211> 23
<212> DNA
<213> Artificial Sequence

<400> 100
tgcagcagtt acttccactc aag 23

<210> 101
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 26 3' RT-PCR primer

<400> 101
aactgaaagc aggattcctc aat 23

<210> 102
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 29 5' RT-PCR primer

<400> 102
ggttaccttg gccgtgagga cac 23

<210> 103
<211> 27
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 29 3' RT-PCR primer

<400> 103
ctctgaactg agaaatgcaa ggagacc 27

<210> 104
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 34 5' RT-PCR primer

<400> 104
ccagcgtctc gcaatgctat cgc 23

<210> 105
<211> 26
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 34 3' RT-PCR primer

<400> 105
cgacatctgg ttcctacttc agggcc 26

<210> 106

<223> Description of Artificial Sequence: synthetic primer: IMX 23 3' RT-PCR primer

<400> 95
atgttgtcca gtctggtctc aag 23

<210> 96
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 24 5' RT-PCR primer

<400> 96
cccagctcaa cgacttcctc cag 23

<210> 97
<211> 25
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 24 3' RT-PCR primer

<400> 97
tgttcatgaa agcggacaca gcagt 25

<210> 98
<211> 25
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 25 5' RT-PCR primer

<400> 98
actgtaatca ggtggcgctt ggttc 25

<210> 99
<211> 28
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 25 3' RT-PCR primer

<400> 99
acctgaattt tcagcataca agggtcac 28

<210> 100
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 26 5' RT-PCR primer

<211> 24
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 35 5' RT-PCR primer

<400> 106
cggccaccat ctttgtacat gtgt 24

<210> 107
<211> 30
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 35 3' RT-PCR primer

<400> 107
ggtccaagat aatcctattt aggctaccaa 30

<210> 108
<211> 24
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 36 5' RT-PCR primer

<400> 108
agagctcatc cagaagggga agga 24

<210> 109
<211> 26
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 36 3' RT-PCR primer

<400> 109
atggtattgg tgattatgtc gccgtt 26

<210> 110
<211> 20
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 37 5' RT-PCR primer

<400> 110
ccggactaca accacgacca 20

<210> 111
<211> 26
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 37 3' RT-PCR primer

<400> 111
ttttgtgttc ttgtagttga aataca 26

<210> 112
<211> 21
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 38 5' RT-PCR primer

<400>112
ccggtgtgga ccatcgagtc a 21

<210> 113
<211> 24
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 38 3' RT-PCR primer

<400> 113
aagccctaga gagtttatta caaa 24

<210> 114
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 45 5' RT-PCR primer

<400> 114
gcggcggcgt tattcccatg atg 23

<210> 115
<211> 27
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 45 3' RT-PCR primer

<400> 115
ggtgctcgct ttatttttca tgtgcaa 27

<210> 116
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 46 5' RT-PCR primer

<400> 116
ccggcgggac gtcagggtct aca 23

<210> 117
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 46 3' RT-PCR primer

<400> 117
gttcccaggc aaccctccct cga 23

<210> 118
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 47 5' RT-PCR primer

<400> 118
ccggccagta cgaaaaggac ttc 23

<210> 119
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 47 3' RT-PCR primer

<400> 119
tagatccata aggctttagt tcc 23

<210> 120
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 48 5' RT-PCR primer

<400> 120
ccggccttaa ggttcctctg aca 23

<210> 121
<211> 24
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 48 3' RT-PCR primer

<400> 121
ggtaggcagc ctgagttgcag ttac 24

<210> 122
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 50 5' RT-PCR primer

<400> 122
catggtcatc actggtctgt ctg 23

<210> 123
<211> 24
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 50 3' RT-PCR primer

<400> 123
cctagaggca tcccagagat agct 24

<210> 124
<211> 23
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 54 5' RT-PCR primer

<400> 124
ctgccctgag accaccatgc tgt 23

<210> 125
<211> 26
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 54 3' RT-PCR primer

<400> 125
gggctacatg tccaaggtgg cttctt 26

<210> 126
<211> 22
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 58 5' RT-PCR primer

<400> 126
ggcagctata gtctggcacc aa 22

<210> 127
<211> 24
<212> DNA
<213> Artificial Sequence

<223> Description of Artificial Sequence: synthetic primer: IMX 58 3' RT-PCR primer

<400> 127
agttctgcag caaagacatg agag 24

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 December 2000 (07.12.2000)

PCT

(10) International Publication Number
WO 00/73324 A3

(51) International Patent Classification[7]: C12Q 1/00, 1/68, C12P 21/06, C12N 1/20, 15/00, A61K 38/00, C07K 1/00, 16/00, C07H 21/02, 21/04

(21) International Application Number: PCT/US00/15191

(22) International Filing Date: 1 June 2000 (01.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/137,058 1 June 1999 (01.06.1999) US

(71) Applicant *(for all designated States except US)*: **DIGITAL GENE TECHNOLOGIES INC.** [US/US]; 11149 North Torrey Pines Road, Suite 110, La Jolla, CA 92037 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **YOUAKIM, Adel** [CA/US]; 7304 175th Street SW, Edmonds, WA 98026 (US). **DUBOSE, Robert, Finley** [US/US]; 6151 156th Place Southeast, Bellevue, WA 98006 (US). **SIMS, John, E.** [US/US]; 4207 43rd Avenue NE, Seattle, WA 98105 (US). **PRIBYL, Thomas, M.** [US/US]; 2609 Pico Place, Apartment 127, San Diego, CA 92109 (US). **HILLBUSH, Brian, S.** [US/US]; 13058 Caminito Angelico, San Diego, CA 92130 (US). **HASEL, Karl, W.** [CH/US]; 247 South Nardo Avenue, Solana Beach, CA 92075 (US).

(74) Agent: **ZIMMERMAN, Roger, P.**; McDonnell Boehnen Hulbert & Berghoff, 300 South Wacker Drive, 32nd Floor, Chicago, IL 60606 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
19 July 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/73324 A3

(54) Title: MODULATION OF GENE EXPRESSION IN GASTROINTESTINAL INFLAMMATION

(57) Abstract: Polynucleotides, polypeptides, kits and methods are provided related to regulated genes characteristic of gastrointestinal inflammation.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/15191

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/4, 6, 69.1, 252.3, 320.1; 530/300, 350, 387.1; 536/23.1, 23.4, 23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

West; stn files - medline, caplus, biosis, biotechds, embase and uspatfull

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MEINHARDT, T. et al. Different genomic structure of mouse and human Lmp7 genes: characterization of MHC-encoded proteasome genes. Immunogenetics. 1993, Vol. 38, pages 373-379, especially page 373 and Figure 1.	1-32
X	FOSS, G. S. Constitutive and interferon-γ-induced expression of the human proteasome subunit multicatalytic endopeptidase complex-like 1. Biochimica et Biophysica Acta. March 1998, Vol. 1402, pages 17-28, especially see Accession No. 13640 for MECL1 cDNA and the MECLI gene in Figure 1 to which Applicants' SEQ ID NO : 15 is 99.4% homologous.	3, 4 & 6

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 14 DECEMBER 2000	Date of mailing of the international search report 26 JAN 2001
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer JOYCE BRIDGERS PARALEGAL SPECIALIST CHEMICAL MATRIX TEKCHAND SAIDHA Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/15191

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CROSS, S. H. Purification of CpG islands using a methylated DNA binding column. Nature Genetics. March 1994, Vol. 6, No. 3, pages 236-244. see accession no. Z65513 has 94% homology to SEQ ID NO : 13.	4 & 6
X	US 5,322,687 A (DONOVAN et al.) 21 June 1994, see SEQ ID NO : 1 and claimed SEQ ID NO : 11.	6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/15191

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/15191

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

C12Q 1/00, 1/68; C12P 21/06; C12N 1/20, 15/00; A61K 38/00; C07K 1/00, 16/00; C07H 21/02, 21/04

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/4, 6, 69.1, 252.3, 320.1; 530/300, 350, 387.1; 536/23.1, 23.4, 23.5

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s)1-15, 19-20, 24-27 & 32, drawn to polynucleotide of SEQ ID NO : 1-10 and the encoded polypeptides, vectors, host cells and related method and kit.

Group II, claim(s)1-15, 19-20, 24-27 & 32, drawn to polynucleotide of SEQ ID NO : 11-20 and the encoded polypeptides, vectors, host cells and related method and kit.

Group III, claim(s)1-15, 19-20, 24-27 & 32, drawn to polynucleotide of SEQ ID NO : 21-30 and the encoded polypeptides, vectors, host cells and related method and kit.

Group IV, claim(s)1-15, 19-20, 24-27 & 32, drawn to polynucleotide of SEQ ID NO : 31-40 and the encoded polypeptides, vectors, host cells and related method and kit.

Group V, claim(s)1-15, 19-20, 24-27 & 32, drawn to polynucleotide of SEQ ID NO : 41-43 and the encoded polypeptides, vectors, host cells and related method and kit.

Group VI, claim(s) 16-18 & 28-31, drawn to antibody of polypeptide of SEQ ID NO : 1-10.

Group VII, claim(s) 16-18 & 28-31, drawn to antibody of polypeptide of SEQ ID NO : 11-20.

Group VIII, claim(s) 16-18 & 28-31, drawn to antibody of polypeptide of SEQ ID NO : 21-30.

Group IX, claim(s) 16-18 & 28-31, drawn to antibody of polypeptide of SEQ ID NO : 31-40.

Group X, claim(s) 16-18 & 28-31, drawn to antibody of polypeptide of SEQ ID NO : 41-43.

Group XI, claim(s) 21-23, drawn to methods of diagnosing, identifying, preventing and treating using the polypeptides or polynucleotides of SEQ ID NOs: 1-10.

Group XII, claim(s) 21-23, drawn to methods of diagnosing, identifying, preventing and treating using the polypeptides or polynucleotides of SEQ ID NOs: 11-20.

Group XIII, claim(s) 21-23, drawn to methods of diagnosing, identifying, preventing and treating using the polypeptides or polynucleotides of SEQ ID NOs: 21-30.

Group XIV, claim(s) 21-23, drawn to methods of diagnosing, identifying, preventing and treating using the polypeptides or polynucleotides of SEQ ID NOs: 31-40.

Group XV, claim(s) 21-23, drawn to methods of diagnosing, identifying, preventing and treating using the polypeptides or polynucleotides of SEQ ID NOs: 41-43.

Unity of Invention - Nucleotide sequences
The Commissioner has decided sua sponte to partially waive 37 CFR 1.475 and 1.499 et seq. to permit applicants to claims up to ten (10) nucleotide sequences which do not have the same or corresponding special technical feature, without the payment of an additional fee. The PCT permits inventions which lack unity of invention to be maintained

in the same international application for payment of additional fees. Thus, in international applications, for each group for which applicant has paid additional international search and/or preliminary examination fees, the PTO has determined that up to four (4) such additional sequences per group is a reasonable number for examination. Further, claims directed to the selected sequences will be examined with claims drawn to any sequence combinations which have a common technical feature with the selected sequences. Nucleotide sequence encoding the same protein are considered to satisfy the unity of invention standard and will continue to be examined together.

The inventions listed as Groups I-XV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Groups I-V have a special technical feature of the nucleotide sequence encoding the polypeptide (see note below), which Groups VI-X and XI-XV do not share. Groups VI-X have a special technical feature of the antibody which Groups I-V and XI-XV do not share. Groups XI-XV have a special technical feature of diagnosis, treatment and prevention, which Groups I-V and Groups VI-X do not share. Thus the various groups discussed above show a lack of unity of invention.

The Commissioner has decided sua sponte to partially waive 37 CFR 1.475 and 1.499 et seq. to permit applicants to claims up to ten (10) nucleotide sequences which do not have the same or corresponding special technical feature, without the payment of an additional fee. The PCT permits inventions which lack unity of invention to be maintained in the same international application for payment of additional fees. Thus, in international applications, for each group for which applicant has paid additional international search and/or preliminary examination fees, the PTO has determined that up to four (4) such additional sequences per group is a reasonable number for examination. Further, claims directed to the selected sequences will be examined with claims drawn to any sequence combinations which have a common technical feature with the selected sequences. Nucleotide sequence encoding the same protein are considered to satisfy the unity of invention standard and will continue to be examined together.